DOCUMENT RESUME

ED 362 539 TM 020 558

AUTHOR	Dossey, John A.; And Others
TITLE	Can Students Do Mathematical Problem Solving? Results from Constructed-Response Questions in NAEP's 1992 Mathematics Assessment.
INSTITUTION	Educational Testing Service, Princeton, NJ. Center for the Assessment of Educational Progress.; National Assessment of Educational Progress, Princeton, NJ.
SPONS AGENCY	National Center for Education Statistics (ED), Washington, DC.
REPORT NO	ETS-R-23-FR01; ISBN-0-16-041921-2; ISBN-0-88685-146-7
PUB DATE	Aug 93
NOTE	237p.
AVAILABLE FROM	U.S. Government Printing Office, Superintendent of Documents, Mail Stop: SSOP, Washington, DC 20402-9328.
PUB TYPE	Statistical Data (110) -- Reports - Descriptive (141)
EDRS PRICE	MF01/PC10 Plus Postage.
DESCRIPTORS	*Constructed Response; *Elementary School Students; Elementary Secondary Education; Grade 4; Grade 8; Grade 12; *Mathematics Tests; Multiple Choice Tests; National Surveys; *Problem Solving; *Secondary School Students; Tables (Data); Test Construction; Testing Programs; *Test Results; Test Use; Thinking Skills
IDENTIFIERS	Large Scale Programs; *National Assessment of Educational Progress

ABSTRACT

In the 1992 mathematics assessment of the National Assessment of Educational Progress (NAEP), approximately one-third of the questions and about half of the student response time were devoted to questions asking students to construct their own responses. Regular constructed response items required a short answer with a problem solution, while extended constructed response required demonstration of student reasoning. Analysis of student papers shows that most made a conscientious effort to respond, but that performance left much to be desired. On regular constructed response questions, the average percent correct by grade level was 42 percent of grade 4, 53 percent of grade 8, and 40 percent of grade 12. For extended constructed response questions, the average percentages of student producing correct responses were 16 percent at grade 4, 8 percent at grade 8, and 9 percent at grade 12. Results indicate that extended constructed response questions can be successfully used in large-scale assessment and that they do signify student proficiency. The levels of performance on multiple-choice items were higher than for constructed responses. Sixty tables present study findings. A procedural appendix explains study methodology, with an additional seven tables. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

NATIONAL CENTER FOR EDUCATION STATISTICS

Can Students Do Mathematical Problem Solving?

Results from Constructed-Response Questions in NAEP's 1992 Mathematics Assessment

ED 362 539

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

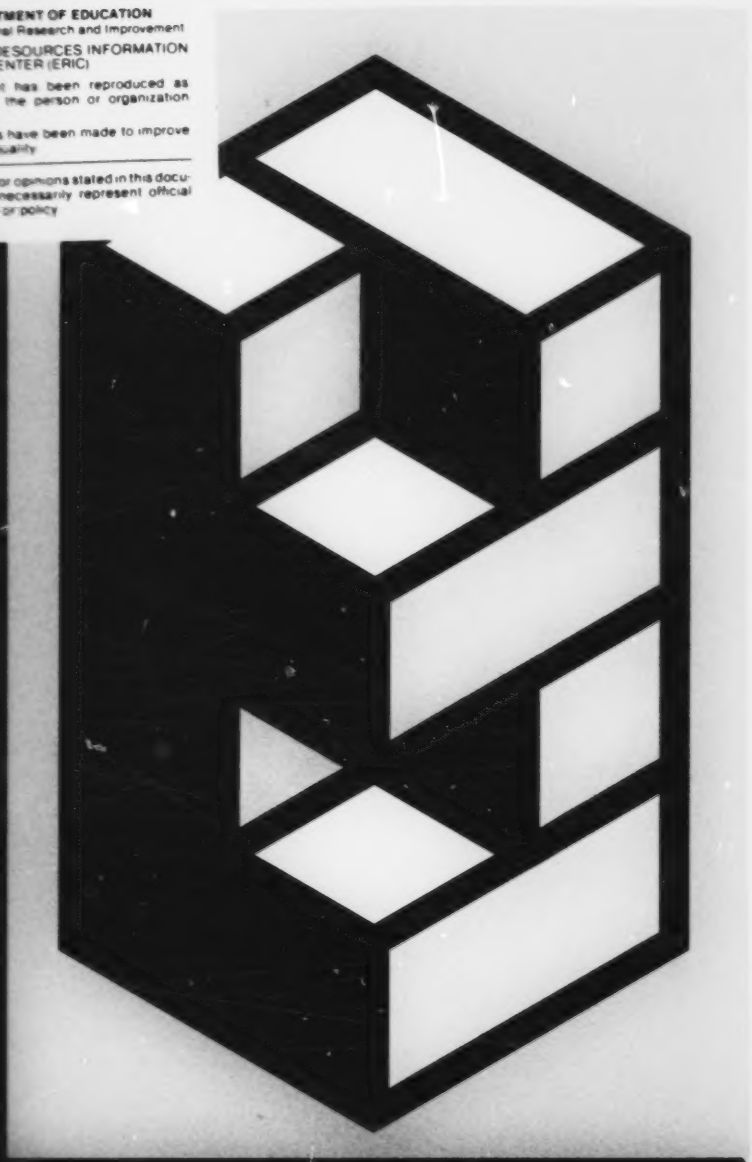
- [x] This document has been reproduced as received from the person or organization originating it.
- [ ] Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

SCOPE OF INTEREST NOTICE

The ERIC Facility has assigned this document for processing to: TM SE

In our judgment, this document is also of interest to the Clearinghouses noted to the right. Indexing should reflect their special points of view.



TM020558

THE NATION'S REPORT CARD

Prepared by Educational Testing Service under contract with the National Center for Education Statistics.

Office of Educational Research and Improvement
U.S. Department of Education

BEST COPY AVAILABLE

What is The Nation's Report Card?

THE NATION'S REPORT CARD, the National Assessment of Educational Progress (NAEP), is the only nationally representative and continuing assessment of what America's students know and can do in various subject areas. Since 1969, assessments have been conducted periodically in reading, mathematics, science, writing, history/geography, and other fields. By making objective information on student performance available to policymakers at the national, state, and local levels, NAEP is an integral part of our nation's evaluation of the condition and progress of education. Only information related to academic achievement is collected under this program. NAEP guarantees the privacy of individual students and their families.

NAEP is a congressionally mandated project of the National Center for Education Statistics, the U.S. Department of Education. The Commissioner of Education Statistics is responsible, by law, for carrying out the NAEP project through competitive awards to qualified organizations. NAEP reports directly to the Commissioner, who is also responsible for providing continuing reviews, including validation studies and solicitation of public comment, on NAEP's conduct and usefulness.

In 1988, Congress created the National Assessment Governing Board (NAGB) to formulate policy guidelines for NAEP. The board is responsible for selecting the subject areas to be assessed, which may include adding to those specified by Congress; identifying appropriate achievement goals for each age and grade; developing assessment objectives; developing test specifications; designing the assessment methodology; developing guidelines and standards for data analysis and for reporting and disseminating results; developing standards and procedures for interstate, regional, and national comparisons; improving the form and use of the National Assessment; and ensuring that all items selected for use in the National Assessment are free from racial, cultural, gender, or regional bias.

The National Assessment Governing Board

Mark D. Musick, Chairman
President
Southern Regional Education Board
Atlanta, Georgia

Hon. William T. Randall, Vice Chair
Commissioner of Education
State Department of Education
Denver, Colorado

Parris C. Battle
Education Specialist
Dade County Public Schools
Miami, Florida

Honorable Evan Bayh
Governor of Indiana
Indianapolis, Indiana

Mary R. Blanton
Attorney
Blanton & Blanton
Salisbury, North Carolina

Boyd W. Boehlje
Attorney and School Board Member
Pella, Iowa

Linda R. Bryant
Dean of Students
Florence Reizenstein Middle School
Pittsburgh, Pennsylvania

Naomi K. Cohen
Office of Policy and Management
State of Connecticut
Hartford, Connecticut

Charlotte Crabtree
Professor
University of California
Los Angeles, California

Chester E. Finn, Jr.
Founding Partner and Senior Scholar
The Edison Project
Washington, DC

Michael S. Glode
Wyoming State Board of Education
Saratoga, Wyoming

William Hume
Chairman of the Board
Basic American, Inc.
San Francisco, California

Christine Johnson
Director of Urban Initiatives
Education Commission of the States
Denver, Colorado

John S. Lindley
Principal
Galloway Elementary School
Henderson, Nevada

Honorable Stephen E. Merrill
Governor of New Hampshire
Concord, New Hampshire

Jason Millman
Professor
Cornell University
Ithaca, New York

Honorable Richard P. Mills
Commissioner of Education
State Department of Education
Montpelier, Vermont

Carl J. Moser
Director of Schools
The Lutheran Church — Missouri Synod
St. Louis, Missouri

John A. Murphy
Superintendent of Schools
Charlotte-Mecklenburg Schools
Charlotte, North Carolina

Michael T. Nettles
Professor
University of Michigan
Ann Arbor, Michigan

Honorable Carolyn Pollan
Arkansas House of Representatives
Fort Smith, Arkansas

Thomas Topuzes
Senior Vice President
Valley Independent Bank
El Centro, California

Marilyn Whirry
English Teacher
Mira Costa High School
Manhattan Beach, California

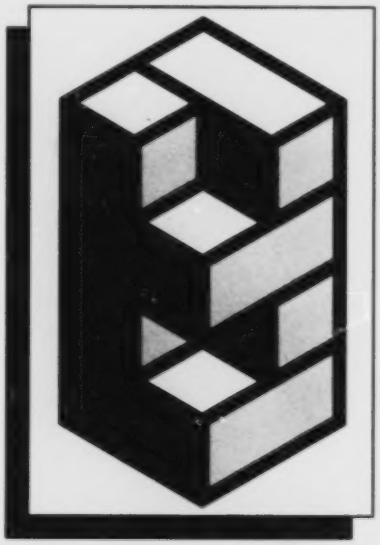
Sharon P. Robinson
Assistant Secretary for Educational
Research and Improvement (Ex-Officio)
U.S. Department of Education
Washington, D.C.

Roy Truby
Executive Director, NAGB
Washington, D.C.

NATIONAL CENTER FOR EDUCATION STATISTICS

Can Students Do Mathematical Problem Solving?

Results from Constructed-Response Questions in NAEP's 1992 Mathematics Assessment



John A. Dossey

Ina V. S. Mullis

Chancey O. Jones

Report No. 23-FR01

August 1993

THE NATION'S REPORT CARD

Prepared by Educational Testing Service under contract with the National Center for Education Statistics

Office of Educational Research and Improvement
U.S. Department of Education

U.S. Department of Education
Richard W. Riley
Secretary

Office of Educational Research and Improvement
Sharon P. Robinson
Assistant Secretary

National Center for Education Statistics
Emerson J. Elliott
Commissioner

FOR MORE INFORMATION:

For ordering information on this report, write:

Education Information Branch
Office of Educational Research and Improvement
U.S. Department of Education
555 New Jersey Avenue, NW
Washington, D.C. 20208-5641

or call 1-800-424-1616 (in the Washington, D.C. metropolitan area call 202-219-1651).

Library of Congress, Catalog Card Number: 93-79253

ISBN: 0-88685-146-7

The work upon which this publication is based was performed for the National Center for Education Statistics, Office of Educational Research and Improvement, by Educational Testing Service.

Educational Testing Service is an equal opportunity, affirmative action employer.

Educational Testing Service, ETS, and *the ETS logo* are registered trademarks of Educational Testing Service.

For sale by the U.S. Government Printing Office
Superintendent of Documents, Mail Stop: SSOP, Washington, DC 20402-9328
ISBN 0-16-041921-2

CONTENTS

EXECUTIVE SUMMARY 1

Major Findings 2

The Scope of NAEP's 1992 Mathematics Assessment 4

INTRODUCTION 5

Figure 1 Description of Mathematics Content Areas 8

Figure 2 Description of Mathematical Abilities 10

Orientation to This Report 11

Table 1 Number of Regular Constructed-Response, Multiple-Choice, and Extended-Response Questions 12

Overview of Assessment Procedures and Methods 14

CHAPTER ONE 17

Performance on Regular Constructed-Response Questions 17

Example Regular Constructed-Response Questions 18

Table 1.1 National Results for Demographic Subgroups for the Regular Constructed-Response Task, "Jill's Class Trip" 20

Table 1.2 State Results for Demographic Subgroups for the Regular Constructed-Response Task, "Jill's Class Trip" 22

Table 1.3 National Results for Demographic Subgroups for the Regular Constructed-Response Task, "Draw a Rectangle on the Grid" 25

Table 1.4 State Results for Demographic Subgroups for the Regular Constructed-Response Task, "Draw a Rectangle on the Grid" 26

Table 1.5 National Results for Demographic Subgroups for the Regular Constructed-Response Task, "Sampling the Yellow and Blue Marbles" 29

Table 1.6 State Results for Demographic Subgroups for the Regular Constructed-Response Task, "Sampling the Yellow and Blue Marbles 30

Table 1.7 National Results for Demographic Subgroups for the Regular Constructed-Response Task, "Find Points on a Grid" 33

Table 1.8 State Results for Demographic Subgroups for the Regular Constructed-Response Task, "Find Points on a Grid" 34

Table 1.9 National Results for Demographic Subgroups for the Regular Constructed-Response Tasks, "Shapes Cut from Cylinders" and "Misleading Pictograph of Trash Cans" 38

Table 1.10 State Results for Demographic Subgroups for the Regular Constructed-Response Tasks, "Shapes Cut from Cylinders" and "Misleading Pictograph of Trash Cans" 39

Table 1.11 National Results for Demographic Subgroups for the Regular Constructed-Response Task, "Two Numbers Make Number Sentence True" 41

Table 1.12 State Results for Demographic Subgroups for the Regular Constructed-Response Task, "Two Numbers Make Number Sentence True" 42

Table 1.13 National Results for Demographic Subgroups for the Regular Constructed-Response Task, "Changing Area of Triangle" 44

Example Constructed-Response Questions Involving Rulers, Protractors, or Manipulative Materials 44

Table 1.14 National Results for Demographic Subgroups for the Regular Constructed-Response Task, "Using Ruler to Measure Centimeters" 47

Table 1.15 State Results for Demographic Subgroups for the Regular Constructed-Response Task, "Using Ruler to Measure Centimeters" 49

Table 1.16 National Results for Demographic Subgroups for the Regular Constructed-Response Task, "Using Ruler to Draw a Rectangle" 54

Table 1.17 State Results for Demographic Subgroups for the Regular Constructed-Response Task, "Using Ruler to Draw a Rectangle" 56

Table 1.18 National Results for Demographic Subgroups for the Regular Constructed-Response Task, "Using Ruler to Draw a Square" 60

Table 1.19 State Results for Demographic Subgroups for the Regular Constructed-Response Task, "Using Ruler to Draw a Square" 61

Table 1.20 National Results for Demographic Subgroups for the Regular Constructed-Response Task, "Using Protractor to Measure an Angle" 64

Table 1.21 State Results for Demographic Subgroups for the Regular Constructed-Response Task, "Using Protractor to Measure an Angle" 65

Example Constructed-Response Questions with Calculator Available 66

Table 1.22 National Results for Demographic Subgroups for the Regular Constructed-Response Tasks, "Estimate Calories" and "Change from Buying Two Calculators" 68

Table 1.23 State Results for Demographic Subgroups for the Regular Constructed-Response Tasks, "Estimate Calories" and "Change from Buying Two Calculators" 69

Table 1.24 National Results for Demographic Subgroups for the Regular Constructed-Response Tasks, "Missing Number in Box" and "Extend Pattern" 72

Table 1.25 State Results for Demographic Subgroups for the Regular Constructed-Response Tasks, "Missing Number in Box" and "Extend Pattern" 73

Table 1.26 National Results for Demographic Subgroups for the Regular Constructed-Response Tasks, "Raymond's Report" and "Area of Trapezoid" 77

Table 1.27 State Results for Demographic Subgroups for the Regular Constructed-Response Tasks, "Raymond's Report" and "Area of Trapezoid" 78

Table 1.28 National Results for Demographic Subgroups for the Regular Constructed-Response Tasks, "Video Rental Costs" and "Area of Parallelogram" 81

Table 1.29 National Results for Demographic Subgroups for the Regular Constructed-Response Tasks, "Graph Pulse Rates" and "F(3.5)" 83

Table 1.30 National Results for Demographic Subgroups for the Regular Constructed-Response Task, "Trigonometry" 85

Summary of National and State Results for Regular Constructed-Response Items 85

CHAPTER TWO 87

Extended Responses to Explain Mathematical Reasoning 87

Introduction 87

Overview of the Tasks and Scoring Guides 87

Figure 2.1 NAEP Mathematics Scoring Category Classifications for Extended Constructed-Response Questions 89

Organization for this Chapter 89

Extended-Response Questions: Grade 4 91

Performance Highlights: Pizza Comparison 96

Table 2.1 National Results for Demographic Subgroups for the Extended-Response Question, "Pizza Comparison" 96

Table 2.2 State Results for Demographic Subgroups for the Extended-Response Question, "Pizza Comparison" 98

Performance Highlights: Laura's Calculator Correction 103

Table 2.3 National Results for Demographic Subgroups for the Extended-Response Question, "Laura's Calculator Correction" 103

Table 2.4 State Results for Demographic Subgroups for the Extended-Response Question, "Laura's Calculator Correction" 105

Performance Highlights: Graphs of Pockets 112

Table 2.5 National Results for Demographic Subgroups for the Extended-Response Question, "Graphs of Pockets" 112

Table 2.6 State Results for Demographic Subgroups for the Extended-Response Question, "Graphs of Pockets" 114

Extended-Response Questions: Grade 8 115

Performance Highlights: Treena's Budget 121

Table 2.7 National Results for Demographic Subgroups for the Extended-Response Task, "Treena's Budget" 121

Table 2.8 State Results for Demographic Subgroups for the Extended-Response Task, "Treena's Budget" 123

Performance Highlights: Radio Stations 129

Table 2.9 National Results for Demographic Subgroups for the Extended-Response Task, "Radio Stations" 129

Table 2.10 State Results for Demographic Subgroups for the Extended-Response Task, "Radio Stations" 131

Performance Highlights: Marcy's Dot Pattern 137

Table 2.11 National Results for Demographic Subgroups for the Extended-Response Task, "Marcy's Dot Pattern" 138

Table 2.12 State Results for Demographic Subgroups for the Extended-Response Task, "Marcy's Dot Pattern" 139

Extended-Response Questions: Grade 12 140

Performance Highlights: Effective Tax Rates 148

Table 2.13 National Results for Demographic Subgroups for the Extended-Response Task, "Effective Tax Rates" 148

Performance Highlights: Patterns of Squares 156

Table 2.14 National Results for Demographic Subgroups for the Extended-Response Task, "Patterns of Squares" 156

Performance Highlights: Graphing Path of Object 163

Table 2.15 National Results for Demographic Subgroups for the Extended-Response Task, "Graphing Path of Object" 163

Summary 165

CHAPTER THREE 167

Summarizing Performance on the Constructed-Response Tasks 167

Difficulty by Question Type and for Subpopulations 167

Table 3.1 National Percentages of Satisfactory or Better Responses to Extended-Response Questions 168

Table 3.2 State Percentages of Satisfactory or Better Responses to Extended-Response Questions 169

Table 3.3 National Percentages of Satisfactory or Better Responses to Extended-Response Questions 170

Table 3.4 State Percentages of Satisfactory or Better Responses to Extended-Response Questions 171

Table 3.5 National Percentages of Satisfactory or Better Responses to Extended-Response Questions 172

Table 3.6 National Average Percentage Correct for Regular Constructed-Response Questions 173

Table 3.7 State Average Percentage Correct for Regular Constructed-Response Questions 175

Table 3.8 National Average Percentage Correct for Multiple-Choice Questions 177

Table 3.9 State Average Percentage Correct for Multiple-Choice Questions 179

The Information Provided by Constructed-Response Tasks 181

FIGURE 3.1 Information Functions 182

Interaction Between Question Type and Curriculum and Instruction 184

Table 3.10 National Average Percentages by Teachers' Reports on the Instructional Emphasis Placed on Reasoning and Communicating 185

Table 3.11 State Average Percentages by Teachers' Reports on the Instructional Emphasis Placed on Reasoning and Communicating 186

Table 3.12 State Average Percentages by Teachers' Reports on the Instructional Emphasis Placed on Reasoning and Communicating 188

Table 3.13 State Average Percentages by Teachers' Reports on the Instructional Emphasis Placed on Reasoning and Communicating 190

Table 3.14 Average Percentages of Successful Responses for Different Question Types by Course Taken in Mathematics 192

Summary 193

Table 3.15 Summary of Average Percentage Correct* by Type of Questions 193

PROCEDURAL APPENDIX 195

Overview of Procedures Used in NAEP's 1992 Mathematics Assessment 197

The Assessment Framework 195

Table A.1 Percentage Distribution of Questions by Grade and Content Area 196

Table A.2 Percentage Distribution of Questions by Grade and Mathematical Ability 196

The Assessment Design 196

National Sampling 198

Table A.3 1992 Student and School Sample Sizes 198

Trial State Assessment Sampling 199

Participation Rates for States and Territories 200

Table A.4 Summary of School and Student Participation 204

Excluded Students 206

Data Collection 207

Scoring 208

Table A.5 Percentages of Exact Agreement for Scoring Reliability Samples for Extended Constructed-Response Tasks 209

Data Analysis and IRT Scaling 209

Per-Item Information Analyses Using IRT 212

Linking the Trial State Results to the National Results 215

NAEP Reporting Groups 215

Table A.6 Percentages of Students in Reporting Groups for the Nation 218

Table A.7 Characteristics of NAEP Students by Race/Ethnicity and by Type of Community 219

Estimating Variability 223

Drawing Inferences from the Results 222

ACKNOWLEDGMENTS 225

EXECUTIVE SUMMARY

As part of the overall education reform initiative, there has been a concerted effort to improve mathematics education in the United States by emphasizing problem solving and application in real-life settings, rather than simply rote memorization. Beginning in the mid-1980s, the National Council of Teachers of Mathematics (NCTM) worked to develop *Curriculum and Evaluation Standards for School Mathematics.* Published in 1989, *The NCTM Standards* emphasize more balanced and dynamic curricular goals where students "do" mathematics -- actively exploring, constructing, and justifying their ideas as they interact and work to solve problems.

The Mathematical Sciences Education Board (MSEB) also has developed a number of publications supporting the need for such reform, stressing the importance of active learning for students, and developing prototypes for assessing mathematics performance in ways that support instructional goals.

There is agreement by the NCTM, MSEB, and a number of national organizations, including The National Council on Education Standards and Testing, that the potential for educational improvement is enhanced if reform is systemic, simultaneously involving such areas as curriculum, instruction, assessments, and professional development. To reinforce reform efforts, assessments should embody the new instructional goals by providing thoughtful problem-solving situations and opportunities for students to explain their approaches.

In NAEP's 1992 mathematics assessment, about one-third of the questions and approximately one-half of the students' response time were devoted to questions asking students to construct their own responses. These questions were classified as regular constructed-response and extended constructed-response tasks. Regular constructed-response questions required students to provide a short answer giving a solution to the problem posed. The extended-response

tasks were a new feature of the 1992 assessment. For these questions, students were allowed at least five minutes for the completion of tasks which required the students to demonstrate -- by writing, by giving examples, or by drawing diagrams -- their mathematical reasoning and problem-solving abilities. For some of the constructed-response questions, NAEP provided students with protractors/rulers, calculators, or "manipulable" geometric shapes.

In general, the analysis of student papers showed that most made a conscientious effort to respond, but the performances exhibited left much to be desired.

MAJOR FINDINGS

- On regular constructed-response questions, which required only a short constructed answer, the average percentage correct by grade level was 42 percent for grade 4, 53 percent for grade 8, and 40 percent for grade 12. Similar performance was noted across the participating states and territories, with the average percentage correct ranging from 27 to 51 percent at grade 4 and from 30 to 63 percent at grade 8. (See Chapter One for examples of specific questions.)

- On extended constructed-response tasks, which required students to solve problems requiring a greater depth of understanding and then explain, at some length, specific features of their solutions, the average percentage of students producing satisfactory or better responses was 16 percent at grade 4, 8 percent at grade 8, and 9 percent at grade 12. Similar performance was noted across the participating states and territories, with the average percentage providing satisfactory or better responses ranging from 7 to 22 percent for grade 4 and from 0 to 13 percent for grade 8. (See Chapter Two for examples of specific tasks.)

- The procedures employed in constructing and scoring extended-response tasks showed that they could be successfully included in a large-scale national assessment and that they significantly contributed to understanding of student proficiency in mathematics at each of the three grades assessed.

    - From approximately one-third to two-thirds of the students provided incorrect responses to these extended questions, indicating little evidence of understanding the mathematics concepts involved or even the question being asked.

- Substantial percentages of students, sometimes as many as one-fifth, simply left their papers blank.

- Most students who did seem to understand the problems had difficulty in explaining their work.

- It is encouraging, however, that some students -- from 1 to 16 percent -- provided *extended* responses to each one of the tasks.

- For the nation and across the states, there was a lower level of performance on both regular and extended constructed-response questions than on the multiple-choice items contained in the 1992 NAEP mathematics assessment.

- For the nation, regardless of question type, there was considerable variation in average performance by students from differing demographic groups at each of the three grades assessed:

  - Average performance for White students was significantly higher than that of Black or Hispanic students.

  - Students in advantaged urban areas performed significantly better than students from disadvantaged urban areas.

  - Students attending private schools performed better than students attending public schools.

- While the extended constructed-response tasks were considerably more difficult than either the multiple-choice or regular constructed-response questions, analyses using item response theory (IRT) showed that this type of question provided considerably more information per item toward understanding student performance for more proficient students than either regular constructed-response or multiple-choice questions.

- Similar analyses showed that the regular constructed-response questions provided more information per item about student proficiency than multiple-choice items for students of any ability and more information about below-average students than the extended constructed-response tasks.

The Scope of NAEP's 1992 Mathematics Assessment

NAEP's 1992 mathematics assessment involved nearly 250,000 fourth-, eighth-, and twelfth-grade students attending approximately 10,000 schools across the nation and the states. The resulting student work, including approximately four million written responses constructed by students in 1992, was scored by professional readers at National Computer Systems in Iowa City, Iowa, using scoring rubrics that had been developed by the NAEP Mathematics Test Development Committee and staff at Educational Testing Service. Each answer to the regular constructed-response questions was scored as receiving credit or not receiving credit. Responses to the extended tasks were evaluated according to a five-point scale ranging from an incorrect to an extended explanation. The scoring rubrics for each question were developed prior to the assessment, revised on the basis of field-test results, and modified a final time following an examination of samples of student responses obtained in the actual assessment. To evaluate the reliability of scoring, 25 percent of the papers for each question were scored by two different scorers. The percentage of exact agreement, averaged across the papers, was 94 percent.

Nationally representative samples of students attending both public and private schools were assessed at grades 4, 8, and 12. In addition, samples of fourth and eighth graders attending public schools were assessed in 44 jurisdictions. NAEP's Trial State Assessment Program in Mathematics was begun in 1990 at grade 8 and expanded in 1992 to include both grades 4 and 8.

In releasing the 1992 mathematics results for the nation and the states, U.S. Education Secretary Richard W. Riley said of the NAEP data that "collectively they mean one thing: hard work, systemic change in all parts of education at the state and local levels, and an increased commitment to a learning ethic in America will all be necessary to move education forward."

INTRODUCTION

The *Curriculum and Evaluation Standards for School Mathematics* developed by the National Council of Teachers of Mathematics (NCTM)[1] have gained wide acceptance in education and in the public arena as a framework for the mathematics that schools should teach.[2] *The NCTM Standards* place particular emphasis on problem solving as central to the curriculum. As described by the NCTM, problem solving "is a primary goal of all mathematics instruction and an integral part of all mathematical activity. Problem solving is not a distinct topic but a process that should permeate the entire program and provide the context in which concepts and skills can be learned."

The NCTM Standards also call for mathematics study to include numerous opportunities for communication, which can be prompted by having students explore, investigate, describe, and explain mathematical ideas through representing, talking, listening, writing, and reading. Underlying the importance of problem solving and communicating mathematically is developing a spirit of inquiry in school mathematics that helps students understand mathematics as reasoning. Because reasoning, communication, and problem solving are central components of successfully doing mathematics, *The NCTM Standards* place these processes at the center of their recommendations for curriculum design and instructional activities.

A mathematics curriculum that fulfills these three seminal standards (which provide a foundation for all others) will differ significantly in both content and instruction from most existing curricula. Consequently, methods for assessing progress toward the vision of the standards also must change. By design, evaluation standards that also emphasize problem solving, communication, and reasoning accompany the NCTM curriculum standards.

---

[1]National Council of Teachers of Mathematics, *Curriculum and Evaluation Standards for School Mathematics* (Reston, VA: National Council of Teachers of Mathematics, 1989).

[2]The National Council on Education Standards and Testing, *Raising Standards for American Education* (Washington, DC: U.S. Department of Education, 1992).

According to the evaluation standards, the following are among assessment practices that should receive more attention:

- Assessing what students know and how they think about mathematics
- Developing problem situations that require the applications of a number of mathematical ideas
- Using calculators, computers, and manipulatives in assessment

As described in *The NCTM Standards*, problem solving in classroom settings requires that the mathematical ideas originate with the students rather than the teacher. The problems to be solved can be from real-world activities, from organized data, and from equations. The strategies to solve them can include using manipulative materials, employing trial and error, making organized lists or tables, drawing diagrams, and looking for patterns. In contrast, it has been observed that commonly used tests continue to stress routine, repetitive, rote tasks instead of offering students opportunities to demonstrate the range of their problem-solving abilities.[3] To help foster improved assessment in mathematics, the Mathematical Sciences Education Board (MSEB) has developed a set of prototype problem-solving tasks for fourth graders that is innovative, challenging, and designed to meet a variety of criteria.[4] For example, the tasks should reflect the "spirit" of the reform movement, promote active mental involvement, emphasize the importance of communicating results rather than isolated answers, allow a variety of creative strategies, and have the potential for influencing instruction positively.

Beginning with the availability of the draft standards and subsequent to their publication in 1989, the National Assessment of Educational Progress (NAEP) has been working toward alignment with *The NCTM Standards*. In 1988, Congress added a new dimension to NAEP by authorizing, on a trial basis,

---

[3] Mathematical Sciences Education Board, *Measuring Up: Prototypes for Mathematics Assessment* (Washington, DC: National Academy Press, 1993).

[4] Mathematical Sciences Education Board, *Measuring Up: Prototypes for Mathematics Assessment* (Washington, DC: National Academy Press, 1993).

voluntary participation in state-level assessments in 1990 and 1992. The program featured mathematics at grade 8 in 1990, and at both grades 4 and 8 in 1992 (as well as at grade 4 in reading in 1992). Because the advent of the Trial State Assessment Program signaled a new era for NAEP, special care was taken to solicit widespread involvement and advice about the development and conduct of the 1990 mathematics assessment. The mathematics objectives framework underlying the assessment, and endorsed for use again in 1992 by the National Assessment Governing Board (NAGB), was developed under the auspices of the Council of Chief State School Officers (CCSSO) through a special NAEP Planning Project sponsored by the National Center for Education Statistics and the National Science Foundation.

The mathematics objectives were designed as a matrix comprising five broad content areas and three levels of mathematical ability. The five content areas are: Numbers and Operations; Measurement; Geometry; Data Analysis, Statistics, and Probability; and Algebra and Functions. The mathematical abilities are Conceptual Understanding, Procedural Knowledge, and Problem Solving.[5] Brief descriptions of the content areas and ability levels follow.

[5] *Mathematics Objectives, 1990 Assessment* (Princeton, NJ: National Assessment of Educational Progress, Educational Testing Service, 1988).

FIGURE 1
Description of Mathematics Content Areas

Numbers and Operations

This content area focuses on students' understanding of numbers (whole numbers, fractions, decimals, and integers) and their application to real-world situations, as well as computational and estimation situations. Understanding numerical relationships as expressed in ratios, proportions, and percents is emphasized. Students' skills in estimation, mental computation, use of calculators, generalization of numerical patterns, and verification of results are also included.

Measurement

This content area focuses on students' ability to describe real-world objects using numbers. Students are asked to identify attributes, select appropriate units, apply measurement concepts, and communicate measurement-related ideas to others. Questions are included that require an ability to read instruments using metric, customary, or nonstandard units with emphasis on precision and accuracy. Questions requiring estimation; measurements; and applications of measurements of length, time, money, temperature, mass/weight, area, volume capacity, and angles are also included under this content area.

Geometry

This content area focuses on students' knowledge of geometric figures and relationships and on their skills in working with this knowledge. These skills are important at all levels of schooling as well as in practical applications. Students need to be able to model and visualize geometric figures in one, two, and three dimensions and to communicate geometric ideas. In addition, students should be able to use informal reasoning to establish geometric relationships.

Data Analysis, Statistics, and Probability

This content area focuses on data representation and analysis across all disciplines and reflects the importance and prevalence of these activities in our society. Statistical knowledge and the ability to interpret data are necessary skills in the contemporary world. Questions emphasize appropriate methods for gathering data, the visual exploration of data, and the development and evaluation of arguments based on data analysis.

Algebra and Functions

This content area is broad in scope, covering a significant portion of the grade 9-12 curriculum, including algebra, elementary functions (pre-calculus), trigonometry, and some topics in discrete mathematics. For the fourth grade, and in part at grade 8, algebraic and functional concepts are treated in more informal, exploratory ways. Proficiency in this content area requires both manipulative facility and conceptual understanding; it involves the ability to use algebra as a means of representation and to use algebraic skills and concepts as problem-solving tools. Functions are viewed not only in terms of algebraic formulas, but also in terms of verbal descriptions, tables of values, and graphs.

* * *

The second feature of the design in the construction of the items was the anticipated cognitive ability required of the student to correctly respond to the question. These three categories, conceptual understanding, procedural knowledge, and problem solving, are detailed in FIGURE 2. The main intent in the use of these categories is to provide balance within each content area among items requiring the use of conceptual knowledge and those requiring procedural skill. The ability category of problem solving requires students to integrate their knowledge of both of the prior areas with their knowledge of problem solving in new situations.

FIGURE 2
Description of Mathematical Abilities

The following three categories of mathematical abilities are not to be construed as hierarchical. For example, problem solving involves interactions between conceptual knowledge and procedural skills, but what is considered complex problem solving at one grade level may be considered conceptual understanding or procedural knowledge at another.

Conceptual Understanding

Students demonstrate conceptual understanding in mathematics when they provide evidence that they can recognize, label, and generate examples and counterexamples of concepts; can use and interrelate models, diagrams and varied representations of concepts; can identify and apply principles; know and can apply facts and definitions; can compare, contrast, and integrate related concepts and principles; can recognize, interpret, and apply the signs, symbols, and terms used to represent concepts; and can interpret the assumptions and relations involving concepts in mathematical settings. Such understandings are essential to performing procedures in a meaningful way and applying them in problem-solving situations.

Procedural Knowledge

Students demonstrate procedural knowledge in mathematics when they provide evidence of their ability to select and apply appropriate procedures correctly, verify and justify the correctness of a procedure using concrete models for symbolic methods, and extend or modify procedures to deal with factors inherent in problem settings. Procedural knowledge includes the various numerical algorithms in mathematics that have been created as tools to meet specific needs in an efficient manner. It also encompasses the abilities to read and produce graphs and tables, execute geometric constructions, and perform noncomputational skills such as rounding and ordering.

Problem Solving

In problem solving, students are required to use their reasoning and analytic abilities when they encounter new situations. Problem solving includes the ability to recognize and formulate problems; determine the sufficiency and consistency of data; use strategies, data, models and relevant mathematics; generate, extend and modify procedures; use reasoning (i.e., spatial, inductive, deductive, statistical and proportional); and judge the reasonableness and correctness of solutions.

The 1987-88 CCSSO project gave special attention to the draft version of *The NCTM Standards* and the nature of formal state objectives and frameworks for mathematics instruction. As a consequence of the major revisions in NAEP's mathematics framework resulting from this effort, Educational Testing Service (ETS) designed new materials and procedures for the 1990 assessment. It included a broad range of questions that required students to solve problems in both constructed-response and multiple-choice formats, provide responses using protractors/rulers, and use calculators (four-function at grade 4 and scientific at grades 8 and 12).

For 1992, to increase NAEP's responsiveness to the then-published standards, the mathematics assessment was nearly doubled in scope to provide greater emphasis on constructed-response questions and innovative problem-solving situations. For 1994, under the direction of NAGB, the NAEP mathematics framework was again modified in light of the standards to ensure continued evolution in future assessments toward the vision espoused by NCTM.[6]

Orientation to This Report

In NAEP's 1992 mathematics assessment, about one-third of the questions and approximately half of the students' response time were devoted to questions asking students to construct their own responses. The apportionment of the 1992 assessment into multiple-choice, regular constructed-response, and extended-response questions is shown in TABLE 1. These questions continued the previous practice of supplying students with protractor/rulers and calculators for portions of the assessment. Also, the assessment was expanded to include "manipulable" geometric shapes. Chapter One of this report presents results for regular constructed-response questions, including those accompanied by tools such as the protractor or ruler, calculator, or geometric shapes. Both national and state-by-state data are provided throughout the report.

Also included for 1992 were extended-response questions which allowed students five minutes or so to demonstrate -- in writing, by giving examples, or by drawing diagrams -- their mathematical reasoning and problem-solving abilities. Five such questions were included at grade 4, and six such questions at grades 8 and 12. Three of these questions at each grade, together with national

---

[6] *1994 National Assessment of Educational Progress: Mathematics Assessment Framework* (Washington, DC: National Assessment Governing Board, U.S. Government Printing Office, 1993).

TABLE 1 Number of Regular Constructed-Response, Multiple-Choice, and Extended-Response† Questions

	Overall	Numbers and Operations	Measurement	Geometry	Data Analysis, Statistics, and Probability	Algebra and Functions
Grade 4						
Regular Constructed-Response	54	20	8	12	9	5
Multiple-Choice	96	41	21	14	10	10
Total: Regular and Multiple-Choice	150	61	29	26	19	15
Extended-Response	5					
Total: Constructed-Response	59					
Grade 8						
Regular Constructed-Response	60	15	12	15	11	7
Multiple-Choice	117	41	19	20	16	21
Total: Regular and Multiple-Choice	177	56	31	35	27	28
Extended-Response	6					
Total: Constructed-Response	66					
Grade 12						
Regular Constructed-Response	58	15	10	10	12	11
Multiple-Choice	115	28	18	21	16	32
Total: Regular and Multiple-Choice	173	43	28	31	28	43
Extended-Response	6					
Total: Constructed-Response	64					

† The extended-response questions are not classified by content area, because they generally cut across domains.

Note: In addition, the national and state assessments included a special assessment in estimation. The multiple-choice questions used in conjunction with the paced audiotape to measure estimation are not included in this table. There were 20 estimation questions at grade 4, and 22 at grades 8 and 12. The counts presented herein reflect the questions included in the analyses discussed in the report.

and state results, are discussed in full in Chapter Two. It should be noted that measuring trends in achievement across time is central to NAEP's purpose. For each assessment, some materials are kept secure and carried forward to future assessments to monitor progress in students' performance. Thus, the example questions presented in this report are, by necessity, those released to the public. Also, because many of these questions were newly introduced in the 1992 assessment, trends are not yet available.

Chapter Three of this report summarizes students' performance on the extended-response questions, the regular constructed-response questions, and the multiple-choice questions, nationally for various demographic groups and for the states. Unless otherwise noted, all differences discussed in this report are statistically significant at the .05 level of significance. This means that the observed differences are unlikely to be due to chance or to sampling variability.

This report is one of a series, which, taken in its entirety, is designed to provide a comprehensive account of the results from NAEP's 1992 mathematics assessment. The *1992 Mathematics Report Card for the Nation and the States* summarizes achievement and compares the results to 1990. Although gains in performance were noted between 1990 and 1992 at all three grade levels assessed, just over 60 percent of the students in grades 4, 8, and 12 were estimated to be at or above the Basic Achievement Level on the 1992 assessment. Across the three grades, 25 percent or fewer were estimated to be at the Proficient Achievement Level or beyond, where students should exhibit evidence of solid academic performance. Most students, particularly at grades 8 and 12, showed success in addition, subtraction, and simple problem solving with whole numbers. Fourth graders had more difficulty solving two-step problems involving multiplication and division. Approximately one-fifth and one-half of the students at grades 8 and 12, respectively, were estimated to have solved problems involving fractions, decimals, and percents as well as elementary concepts in geometry, statistics, and algebra. The low levels of performance described in *The 1992 Mathematics Report Card* invite a more detailed look at student performance, particularly on the sorts of tasks presented in Chapter Two of this report.

Other reports of findings from NAEP's 1992 mathematics assessment, including those specifically tailored for each participating state, highlight performance results and relationships between achievement and background factors. The *Data Compendium for the NAEP 1992 Mathematics Assessment of the Nation and the States* provides a detailed resource of assessment results.

Overview of Assessment Procedures and Methods

NAEP's 1992 mathematics assessment included nearly 250,000 fourth-, eighth-, and twelfth-grade students attending approximately 10,000 schools across the nation and the states. Nationally representative samples of students attending both public and private schools were assessed at grades 4, 8, and 12. Additionally, samples of fourth and eighth graders attending public schools were assessed in 44 jurisdictions.

These participants include:

Alabama	**Louisiana**	**Ohio**
Arizona	**Maine**	**Oklahoma**
Arkansas	**Maryland**	**Pennsylvania**
California	**Massachusetts**	**Rhode Island**
Colorado	**Michigan**	**South Carolina**
Connecticut	**Minnesota**	**Tennessee**
Delaware	**Mississippi**	**Texas**
District of Columbia	**Missouri**	**Utah**
Florida	**Nebraska**	**Virginia**
Georgia	**New Hampshire**	**West Virginia**
Hawaii	**New Jersey**	**Wisconsin**
Idaho	**New Mexico**	**Wyoming**
Indiana	**New York**	
Iowa	**North Carolina**	**Guam**
Kentucky	**North Dakota**	**Virgin Islands***

*The Virgin Islands participated in the testing portion of The 1992 Trial State Assessment Program. However, in accordance with the legislation providing for participants to review and give permission for release of their results, the Virgin Islands chose not to release their results at grade 4 in the 1992 NAEP reports.

All NAEP data are collected by trained administrators. Data for the national assessment were collected by a field staff managed by the ETS subcontractor, Westat, Inc. However, in accordance with the NAEP legislation, data collection for the Trial State Assessment Program was the responsibility of each participating jurisdiction. Uniformity of procedures across states was achieved through training and quality control monitoring by Westat, Inc. Westat staff trained nearly 10,000 state assessment administrators using a video presentation accompanied by a scripted trainer's guide and practice exercises.

BEST COPY AVAILABLE

Quality control was provided by unannounced, random monitoring of half the sessions in each state. The results of the monitoring indicated a high degree of quality and uniformity across sessions.

The participation rates for the nation and the states are found in the Procedural Appendix (in particular, see TABLE A.4). It should be noted that several states did not satisfy the guidelines for participation rates. Further analyses, documented in the *Technical Report of the 1992 Trial State Assessment in Mathematics*, suggest that nonresponse bias, if any, was probably quite small. Nevertheless, Delaware, Maine, Nebraska, New Hampshire, New Jersey, New York, Oklahoma, and Guam received notations at grade 4 and Alabama, Maine, Nebraska, New Jersey, and New York received notations at grade 8 for not meeting one or more of the guidelines for participation rates.

The materials, including approximately four million written responses constructed by students in 1992, were scored by a second subcontractor, National Computer Systems, and the results were analyzed by Educational Testing Service. As expected, numerous quality control steps were undertaken to ensure the accuracy of the results. Throughout, NCES and its contractors worked closely with the Trial State Assessment NETWORK, which includes representatives from all interested states. Federal funding permitted state education personnel to meet with staff members from NCES, the contractors, NAGB, and CCSSO at NETWORK meetings regularly held to review NAEP materials and procedures. Further details about the methods and procedures used in NAEP's 1992 mathematics assessment of the nation and states are provided in the Procedural Appendix and the *Technical Report of the 1992 Trial State Assessment in Mathematics*.

CHAPTER ONE

Performance on Regular Constructed-Response Questions

The 1992 NAEP assessments at grades 4, 8, and 12 required students to supply their own written responses, in one form or another, to more than one-third of the questions. Most of these questions were classified as "regular" constructed-response questions. These questions asked students to carry out a calculation and write an answer; to examine a situation and describe why one alternative or another was correct; or to measure or draw a geometric figure given some boundary conditions. Profiles of student achievement on these questions, when combined with information from the extended-response questions discussed in Chapter Two of this report, provide a broader view of students' mathematical abilities than that possible from multiple-choice items alone. Information in Chapter Three provides an examination of student proficiency on the constructed-response questions relative to their performance on other types of items in the assessment.

While the regular constructed-response questions do not demand extensive amounts of student investigation or ask students to show their work, they do move the assessment of students' content knowledge beyond the selection of a response from a list, as in the multiple-choice format. The constructed-response questions examined in this chapter are of three varieties. The first includes questions where students had access to neither manipulative materials nor to a calculator. The second comprises those where students had access to rulers, protractors, or manipulative materials. The third consists of questions where students had access to calculators in providing their responses. This chapter presents samples of student work on these three types of questions together with results at the national and state levels. National results are provided for demographic subgroups.

The constructed-response questions were scored by professional readers who had experience in education. These readers were thoroughly trained and subsequently worked to evaluate the 4 million student-constructed responses collected as part of NAEP's 1992 mathematics assessment. The scoring was conducted at National Computer Systems in Iowa City, Iowa, using rubrics that had been developed by the NAEP Mathematics Test Development Committee and

ETS staff. Each answer to the regular constructed-response questions was scored either as receiving credit or not receiving credit. Responses to the extended questions were evaluated according to a five-point scale ranging from incorrect to extended (see Chapter Two). The scoring rubrics were developed prior to the assessment, revised on the basis of field-test results, and modified a final time following an examination of samples of student responses obtained from the actual assessment. To determine the reliability of the scoring, 25 percent of the students' responses to each question were evaluated by two different scorers. The percentage of exact agreement between readers, averaged across questions for both the national and Trial State Assessment reliability samples, was 94 percent (see Procedural Appendix for further details).

Example Regular Constructed-Response Questions

The following eight examples illustrate student performance on regular constructed-response questions, for which the students had access to neither a calculator nor to manipulative materials. The questions presented throughout this report are representative of the totals of 59, 66, and 64 constructed-response questions included in the overall assessment at grades 4, 8, and 12, respectively. Other examples of constructed-response questions released to the public can be found in the *Data Compendium for the NAEP 1992 Mathematics Assessment of the Nation and the States.*[7]

The first question shown below is from the Numbers and Operations content area and was included in the assessments at both grades 4 and 8. At grade 4, 22 percent of the students correctly responded to the item, while at grade 8, 59 percent of the students correctly stated that Jill would have to work three weeks in order to earn the amount of money needed for the class trip. Considering that the problem requires little more than finding the weekly total earned and comparing this total with the $45 needed, the somewhat low level of performance at both grades provides insight into students' difficulty with numbers and operations questions calling for more than one step to the solution.

[7] *Data Compendium for the NAEP 1992 Mathematics Assessment of the Nation and the States* (Washington, DC: National Center for Education Statistics, U.S. Government Printing Office, 1993).

EXAMPLE 1: Numbers and Operations

Overall Percent Correct*
Grade 4 -- 22 (1.4)
Grade 8 -- 59 (1.3)

Jill needs to earn $45.00 for a class trip. She earns $2.00 each day on Mondays, Tuesdays, and Wednesdays, and $3.00 each day on Thursdays, Fridays, and Saturdays. She does not work on Sundays. How many weeks will it take her to earn $45.00 ?

Answer: 3 weeks

*The standard errors of the estimated percentages appear in parentheses.

TABLE 1.1 provides data giving the percentage of correct responses to this question for the nation and demographic subpopulations. The results provided in this and other corresponding national tables are based on students attending both private and public schools. Also, it should be noted that the data for all students, regardless of whether their racial/ethnic group is reported separately, were included in computing the overall results. Assessment data for Asian/Pacific Islander and American Indian students are not reported separately in the tables containing national results because there were too few participating students in those classifications to permit stable results for individual questions.[8] Definitions and distributions of subpopulations as well as discussions of sampling and analysis procedures are found in the Procedural Appendix.

Some differences in performance can be seen among subpopulations. In general, these follow patterns seen in NAEP and other educational achievement data.[9] For example, at both grades 4 and 8, White students performed better on this question than did their Black and Hispanic counterparts, and students attending schools in advantaged urban communities had higher percentages of correct responses than did students attending schools in disadvantaged urban communities.

---

[8] The sample sizes responding to individual questions were approximately 1,600 students per grade for the nation and 650 students per grade for the states. These sample sizes are in contrast to the total sample sizes of approximately 9,000 students per grade for the nation and 2,500 students per state that form the bases of the aggregated results across questions provided in most NAEP reports. The aggregated data permit reporting for more student subgroups than is possible for individual questions.

[9] Ina V.S. Mullis, John A. Dossey, Eugene H. Owen, and Gary W. Phillips, *NAEP 1992 Mathematics Report Card for the Nation and the States* (Washington, DC: National Center for Education Statistics, U.S. Government Printing Office, 1993).

TABLE 1.1 National Results for Demographic Subgroups for the Regular Constructed-Response Task, "Jill's Class Trip"

Grade 4

	Correct	Incorrect	No Response
Nation	22 (1.4)	70 (1.5)	8 (0.9)
Northeast	28 (3.9)	66 (3.4)	7 (1.5)
Southeast	16 (3.0)	77 (3.7)	7 (1.9)
Central	22 (2.4)	70 (2.4)	7 (2.0)
West	23 (2.6)	66 (2.5)	11 (1.8)
White	26 (1.8)	66 (1.8)	8 (1.0)
Black	11 (2.8)	80 (3.1)	8 (2.1)
Hispanic	13 (2.1)	77 (3.1)	10 (2.1)
Male	21 (1.7)	69 (1.8)	10 (1.3)
Female	23 (2.0)	71 (2.1)	6 (1.1)
Advantaged Urban	34 (3.7)	60 (3.5)	6 (1.8)
Disadvantaged Urban	10 (2.0)	76 (3.8)	14 (3.7)
Extreme Rural	19 (3.0)	67 (3.5)	14 (3.2)
Other	22 (1.7)	71 (1.8)	6 (0.9)
Public	22 (1.6)	70 (1.7)	8 (1.0)
Catholic and Other Private	23 (2.4)	70 (2.7)	7 (1.7)

Grade 8

	Correct	Incorrect	No Response
Nation	59 (1.3)	38 (1.2)	4 (0.4)
Northeast	59 (2.7)	38 (2.3)	4 (0.9)
Southeast	53 (2.7)	43 (2.2)	4 (1.0)
Central	63 (2.6)	33 (2.3)	4 (1.0)
West	61 (2.8)	37 (3.0)	3 (0.6)
White	65 (1.6)	33 (1.6)	2 (0.4)
Black	37 (3.6)	53 (3.2)	10 (2.0)
Hispanic	51 (3.2)	46 (3.3)	4 (0.9)
Male	56 (1.6)	40 (1.6)	4 (0.8)
Female	62 (1.7)	35 (1.7)	3 (0.7)
Advantaged Urban	67 (5.2)	32 (4.8)	1 (0.5)
Disadvantaged Urban	42 (4.8)	43 (4.5)	15 (2.6)
Extreme Rural	62 (5.3)	37 (5.1)	2 (0.5)
Other	59 (1.8)	38 (1.6)	3 (0.5)
Public	58 (1.4)	38 (1.3)	4 (0.5)
Catholic and Other Private	66 (2.9)	32 (2.7)	2 (0.7)

The standard errors of the estimated percentages appear in parentheses. It can be said with about 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details). Percentages may not total 100 percent due to rounding error.

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment

The state-by-state data in TABLE 1.2 for students in grades 4 and 8 show considerable variation. For the fourth graders, performance ranged from 12 to 28 percent correct. At grade 8, student performance was higher, but still extremely variable, ranging from 37 to 71 percent correct. It should be noted that the regional results shown in the state tables are based on the nationally and regionally representative samples of public-school students who were assessed as part of the national program, and not from an aggregate of the separate state-by-state tables. The assignment of states to the four regions is described in the Procedural Appendix. Using the regional results from the national program is necessary because the voluntary nature of the Trial State Program did not guarantee representative regional results from the aggregated data across states, since not all states participated. Also, because the state assessment results are based only on students attending public schools, the regional results in the state tables (also based only on public-school students) should be used in making comparisons between national, regional, and state performance.

TABLE 1.2 | **Percentage of Correct Responses to Regular Constructed-Response Question, "Jill's Class Trip"**

PUBLIC SCHOOLS	Grade 4 - 1992		
	Correct	Incorrect	No Response
NATION	22 (1.6)	70 (1.7)	8 (1.0)
Northeast	27 (4.4)	66 (3.8)	6 (1.5)
Southeast	16 (3.4)	77 (4.2)	7 (2.1)
Central	23 (2.8)	70 (3.0)	7 (2.1)
West	23 (2.9)	66 (2.5)	11 (2.0)
STATES			
Alabama	17 (1.6)	77 (1.7)	6 (1.0)
Arizona	19 (1.5)	74 (1.8)	7 (1.0)
Arkansas	17 (1.5)	79 (1.7)	4 (0.8)
California	20 (1.7)	71 (2.0)	8 (1.4)
Colorado	20 (1.8)	73 (2.1)	7 (0.8)
Connecticut	27 (2.0)	64 (2.4)	8 (1.2)
Delaware	19 (2.0)	76 (2.2)	5 (0.9)
Dist. Columbia	12 (1.2)	81 (1.5)	7 (1.1)
Florida	15 (1.6)	79 (1.7)	6 (0.8)
Georgia	19 (1.6)	76 (1.5)	4 (0.8)
Hawaii	23 (1.7)	71 (1.9)	7 (1.4)
Idaho	18 (1.7)	74 (1.8)	7 (1.0)
Indiana	21 (2.0)	75 (2.0)	3 (0.7)
Iowa	25 (2.0)	71 (2.0)	3 (0.6)
Kentucky	21 (2.0)	76 (2.0)	4 (0.8)
Louisiana	14 (1.7)	80 (1.8)	6 (1.0)
Maine	25 (1.9)	72 (1.9)	3 (0.9)
Maryland	24 (1.6)	70 (1.7)	6 (0.9)
Massachusetts	25 (2.3)	70 (2.5)	5 (1.0)
Michigan	21 (1.7)	74 (1.8)	6 (0.9)
Minnesota	28 (1.7)	67 (1.8)	5 (1.1)
Mississippi	13 (1.6)	81 (1.6)	6 (0.9)
Missouri	21 (2.0)	74 (2.0)	4 (0.9)
Nebraska	26 (2.2)	70 (2.1)	5 (1.2)
New Hampshire	25 (1.8)	67 (2.0)	8 (1.2)
New Jersey	23 (2.1)	69 (2.0)	8 (1.2)
New Mexico	17 (1.5)	78 (1.9)	6 (1.0)
New York	24 (1.8)	73 (2.0)	3 (0.7)
North Carolina	17 (1.7)	78 (1.8)	5 (0.8)
North Dakota	24 (1.8)	72 (2.0)	4 (0.9)
Ohio	22 (1.7)	73 (1.8)	5 (0.8)
Oklahoma	23 (1.8)	73 (2.0)	4 (0.8)
Pennsylvania	24 (1.8)	72 (1.8)	3 (0.7)
Rhode Island	17 (2.0)	77 (2.0)	6 (1.0)
South Carolina	15 (1.6)	80 (1.9)	5 (0.8)
Tennessee	19 (1.9)	76 (2.0)	5 (1.0)
Texas	19 (1.7)	77 (1.7)	3 (0.8)
Utah	22 (1.5)	72 (1.7)	6 (0.9)
Virginia	23 (2.0)	72 (2.2)	5 (0.9)
West Virginia	15 (1.2)	79 (1.4)	6 (0.9)
Wisconsin	23 (1.8)	72 (2.0)	4 (0.9)
Wyoming	20 (1.6)	75 (1.9)	6 (0.9)
TERRITORY			
Guam	12 (1.6)	81 (1.9)	7 (1.3)

The standard errors of the estimated percentages appear in parentheses. It can be said with 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details). Percentages may not total 100 percent due to rounding error.

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment.

BEST COPY AVAILABLE

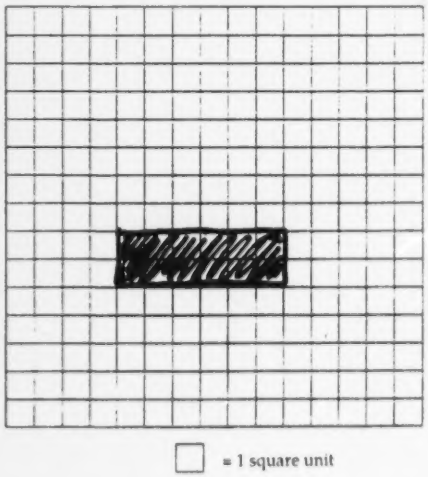
TABLE 1.2 | **Percentage of Correct Responses to Regular Constructed-Response Question, "Jill's Class Trip" (continued)**

PUBLIC SCHOOLS	Grade 8 - 1992		
	Correct	Incorrect	No Response
NATION	58 (1.4)	39 (1.3)	4 (0.5)
Northeast	58 (2.7)	39 (2.3)	4 (1.1)
Southeast	52 (2.9)	44 (2.4)	5 (1.1)
Central	62 (2.7)	34 (2.5)	4 (1.1)
West	60 (2.9)	37 (3.1)	3 (0.7)
STATES			
Alabama	55 (2.3)	42 (2.3)	3 (0.8)
Arizona	65 (2.1)	32 (1.9)	3 (0.8)
Arkansas	60 (2.5)	38 (2.3)	3 (0.8)
California	61 (1.9)	35 (1.9)	4 (0.7)
Colorado	63 (1.9)	35 (2.0)	2 (0.6)
Connecticut	67 (2.0)	32 (2.2)	1 (0.4)
Delaware	59 (2.5)	39 (2.5)	3 (0.9)
Dist. Columbia	46 (2.3)	49 (2.6)	5 (0.9)
Florida	57 (2.4)	38 (2.4)	5 (1.1)
Georgia	58 (2.1)	39 (2.2)	3 (0.8)
Hawaii	55 (2.2)	39 (2.2)	7 (1.0)
Idaho	68 (1.9)	29 (2.0)	3 (0.7)
Indiana	61 (2.0)	37 (1.8)	2 (0.5)
Iowa	71 (2.3)	28 (2.1)	1 (0.5)
Kentucky	61 (1.9)	36 (2.0)	2 (0.7)
Louisiana	54 (2.3)	43 (2.1)	3 (0.7)
Maine	69 (2.4)	30 (2.3)	1 (0.4)
Maryland	59 (2.2)	36 (2.4)	5 (0.9)
Massachusetts	63 (1.8)	34 (1.8)	3 (0.7)
Michigan	63 (1.8)	35 (1.8)	1 (0.6)
Minnesota	68 (1.8)	31 (1.8)	1 (0.3)
Mississippi	51 (2.1)	44 (1.9)	4 (0.8)
Missouri	62 (2.1)	36 (2.0)	2 (0.6)
Nebraska	65 (2.6)	34 (2.6)	1 (0.3)
New Hampshire	67 (2.1)	31 (2.1)	2 (0.6)
New Jersey	68 (1.8)	31 (1.8)	1 (0.5)
New Mexico	56 (1.8)	40 (1.9)	4 (0.7)
New York	63 (2.4)	34 (2.3)	3 (0.7)
North Carolina	59 (2.0)	39 (2.0)	2 (0.5)
North Dakota	70 (1.8)	29 (1.9)	2 (0.7)
Ohio	65 (2.1)	33 (2.1)	2 (0.6)
Oklahoma	64 (2.0)	32 (2.2)	4 (0.9)
Pennsylvania	64 (2.1)	34 (2.0)	3 (0.8)
Rhode Island	61 (1.8)	36 (1.8)	3 (0.9)
South Carolina	61 (2.0)	38 (2.0)	1 (0.5)
Tennessee	57 (2.2)	40 (2.4)	3 (0.8)
Texas	58 (2.3)	38 (2.1)	4 (0.8)
Utah	68 (1.9)	30 (1.8)	2 (0.6)
Virginia	64 (2.2)	35 (2.1)	1 (0.4)
West Virginia	59 (2.1)	37 (2.1)	4 (0.8)
Wisconsin	66 (2.5)	31 (2.3)	2 (0.6)
Wyoming	65 (2.2)	32 (2.1)	3 (0.7)
TERRITORIES			
Guam	44 (2.8)	49 (2.3)	7 (1.6)
Virgin Islands	37 (2.8)	51 (2.8)	12 (1.6)

The following regular constructed-response question asking students to draw a rectangle having an area of 12 on the provided grid was classified in the Measurement content area. Students were required to draw any one of several possible responses, including rectangles of dimensions 1 by 12, 2 by 6, or 3 by 4. (No student drew a rectangle involving rational number dimensions.) Percentages of correct responses on this task, also given to students at both grades 4 and 8, were higher than the previous item, particularly at grade 4, with 42 percent of the fourth graders and 66 percent of the eighth graders answering correctly. State performance varied from 24 to 54 percent correct for grade 4 students and from 38 to 78 percent correct for grade 8 students.

EXAMPLE 2: Measurement

On the grid below, draw a rectangle with an area of 12 square units.



= 1 square unit

Overall Percent Correct*
Grade 4 -- 42 (1.4)
Grade 8 -- 66 (1.5)

(One of the possible answers)

*The standard errors of the estimated percentages appear in parentheses.

TABLE 1.3 National Results for Demographic Subgroups for the Regular Constructed-Response Task, "Draw a Rectangle on the Grid"

Grade 4

	Correct	Incorrect	No Response
Nation	42 (1.4)	50 (1.4)	7 (0.8)
Northeast	50 (2.9)	43 (2.1)	8 (2.1)
Southeast	38 (3.5)	54 (3.0)	8 (1.4)
Central	45 (3.0)	50 (3.5)	6 (1.8)
West	39 (2.1)	53 (1.9)	8 (1.7)
White	49 (1.9)	46 (1.8)	5 (0.8)
Black	22 (2.6)	64 (2.4)	15 (2.5)
Hispanic	30 (2.4)	58 (2.7)	12 (2.1)
Male	43 (2.0)	49 (2.1)	9 (1.2)
Female	42 (2.0)	52 (1.9)	6 (1.0)
Advantaged Urban	54 (3.8)	42 (3.3)	4 (1.6)
Disadvantaged Urban	24 (3.6)	56 (3.8)	20 (4.5)
Extreme Rural	39 (4.8)	54 (4.7)	6 (2.0)
Other	44 (1.8)	50 (1.7)	6 (0.9)
Public	43 (1.5)	50 (1.5)	8 (0.9)
Catholic and Other Private	41 (3.6)	53 (3.3)	6 (1.7)

Grade 8

	Correct	Incorrect	No Response
Nation	66 (1.5)	31 (1.3)	3 (0.5)
Northeast	64 (4.3)	34 (4.2)	2 (0.4)
Southeast	64 (2.6)	33 (1.8)	3 (1.4)
Central	69 (2.3)	29 (2.2)	2 (0.9)
West	69 (2.7)	28 (2.4)	3 (0.8)
White	72 (1.9)	26 (1.8)	1 (0.4)
Black	46 (3.5)	48 (3.2)	6 (1.9)
Hispanic	56 (3.4)	38 (2.9)	6 (1.8)
Male	67 (2.1)	30 (1.9)	3 (0.8)
Female	66 (1.4)	32 (1.5)	2 (0.5)
Advantaged Urban	74 (4.0)	24 (4.0)	2 (1.2)
Disadvantaged Urban	51 (4.4)	44 (4.9)	5 (1.4)
Extreme Rural	61 (3.6)	37 (3.5)	2 (0.8)
Other	68 (1.7)	30 (1.5)	3 (0.6)
Public	66 (1.6)	32 (1.5)	3 (0.6)
Catholic and Other Private	72 (2.5)	26 (2.6)	2 (0.6)

The standard errors of the estimated percentages appear in parentheses. It can be said with about 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details). Percentages may not total 100 percent due to rounding error.

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment

TABLE 1.4 | **Percentage of Correct Responses to Regular Constructed-Response Question, "Draw a Rectangle on the Grid"**

PUBLIC SCHOOLS	Grade 4 - 1992		
	Correct	Incorrect	No Response
NATION	43 (1.5)	50 (1.5)	8 (0.9)
Northeast	50 (3.1)	42 (2.2)	8 (2.5)
Southeast	37 (3.9)	54 (3.3)	9 (1.6)
Central	46 (3.2)	48 (3.8)	5 (1.8)
West	40 (2.3)	53 (2.2)	8 (1.6)
STATES			
Alabama	34 (2.4)	59 (2.6)	8 (1.2)
Arizona	37 (1.7)	57 (1.7)	7 (1.1)
Arkansas	35 (2.1)	59 (2.2)	6 (1.0)
California	39 (2.3)	53 (2.3)	8 (1.3)
Colorado	47 (2.2)	47 (2.2)	6 (1.1)
Connecticut	54 (2.2)	41 (2.1)	6 (1.0)
Delaware	39 (2.4)	53 (2.6)	8 (1.4)
Dist. Columbia	28 (1.8)	58 (2.0)	14 (1.3)
Florida	41 (2.6)	52 (2.8)	7 (1.1)
Georgia	40 (2.4)	54 (2.0)	6 (0.9)
Hawaii	47 (2.4)	48 (2.3)	5 (1.0)
Idaho	45 (2.2)	49 (2.3)	7 (1.2)
Indiana	48 (2.5)	48 (2.3)	5 (0.8)
Iowa	50 (2.3)	46 (2.0)	4 (0.8)
Kentucky	39 (2.5)	56 (2.3)	5 (0.8)
Louisiana	33 (2.1)	58 (2.2)	9 (1.4)
Maine	52 (2.4)	46 (2.3)	2 (0.7)
Maryland	46 (1.8)	48 (1.9)	6 (1.2)
Massachusetts	44 (2.4)	48 (2.5)	8 (1.2)
Michigan	41 (2.2)	56 (2.1)	3 (0.7)
Minnesota	52 (2.7)	43 (2.6)	5 (0.9)
Mississippi	29 (1.9)	60 (2.1)	11 (1.7)
Missouri	41 (2.3)	56 (2.3)	3 (0.6)
Nebraska	45 (2.6)	51 (2.6)	4 (0.8)
New Hampshire	49 (2.3)	45 (2.3)	6 (1.0)
New Jersey	47 (2.5)	46 (2.3)	7 (1.2)
New Mexico	36 (2.6)	58 (2.7)	7 (1.4)
New York	40 (2.2)	53 (2.5)	8 (1.2)
North Carolina	38 (2.1)	55 (2.3)	7 (0.9)
North Dakota	43 (2.8)	54 (2.5)	4 (0.9)
Ohio	49 (2.7)	46 (2.6)	5 (1.0)
Oklahoma	37 (2.4)	58 (2.5)	5 (1.1)
Pennsylvania	49 (2.3)	47 (2.2)	4 (0.6)
Rhode Island	38 (2.3)	53 (2.1)	9 (1.2)
South Carolina	39 (2.2)	54 (2.1)	7 (1.0)
Tennessee	37 (2.1)	53 (2.2)	10 (1.1)
Texas	47 (2.9)	48 (2.6)	5 (1.0)
Utah	49 (2.3)	46 (2.3)	5 (1.0)
Virginia	40 (2.3)	55 (2.2)	5 (1.1)
West Virginia	38 (2.2)	57 (2.2)	5 (1.0)
Wisconsin	45 (2.0)	51 (2.1)	5 (0.8)
Wyoming	46 (2.2)	50 (2.4)	4 (0.9)
TERRITORY			
Guam	24 (2.1)	61 (2.4)	15 (1.7)

The standard errors of the estimated percentages appear in parentheses. It can be said with 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details). Percentages may not total 100 percent due to rounding error.

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment.

TABLE 1.4 | **Percentage of Correct Responses to Regular Constructed-Response Question, "Draw a Rectangle on the Grid" (continued)**

PUBLIC SCHOOLS	Grade 8 - 1992		
	Correct	Incorrect	No Response
NATION	66 (1.6)	32 (1.5)	3 (0.6)
Northeast	63 (5.2)	35 (5.1)	2 (0.5)
Southeast	62 (2.8)	34 (1.9)	4 (1.6)
Central	68 (2.5)	30 (2.4)	2 (1.0)
West	69 (2.8)	28 (2.5)	3 (0.9)
STATES			
Alabama	54 (2.6)	42 (2.4)	4 (1.0)
Arizona	64 (2.4)	33 (2.3)	3 (0.7)
Arkansas	61 (1.8)	37 (1.9)	2 (0.5)
California	64 (2.3)	32 (2.3)	4 (0.9)
Colorado	69 (2.2)	28 (2.0)	3 (0.7)
Connecticut	71 (1.8)	28 (1.8)	1 (0.7)
Delaware	60 (2.7)	36 (2.7)	3 (0.9)
Dist. Columbia	49 (2.4)	43 (2.1)	8 (1.5)
Florida	58 (2.1)	37 (2.2)	4 (0.9)
Georgia	65 (1.8)	33 (1.7)	2 (0.7)
Hawaii	62 (2.3)	33 (2.1)	5 (1.0)
Idaho	67 (1.8)	31 (1.8)	2 (0.5)
Indiana	62 (2.5)	36 (2.7)	2 (0.7)
Iowa	78 (2.0)	21 (1.9)	1 (0.4)
Kentucky	61 (1.8)	36 (1.7)	3 (0.6)
Louisiana	57 (2.2)	38 (2.0)	5 (0.8)
Maine	72 (1.9)	27 (1.9)	1 (0.5)
Maryland	61 (2.5)	35 (2.4)	4 (0.8)
Massachusetts	63 (2.5)	34 (2.5)	3 (1.0)
Michigan	66 (2.3)	31 (2.2)	3 (0.9)
Minnesota	76 (1.7)	22 (1.6)	1 (0.5)
Mississippi	51 (2.4)	44 (2.2)	5 (1.2)
Missouri	69 (1.8)	29 (1.8)	2 (0.6)
Nebraska	68 (2.5)	31 (2.4)	2 (0.5)
New Hampshire	72 (1.9)	27 (2.0)	1 (0.5)
New Jersey	69 (2.5)	28 (2.4)	3 (0.9)
New Mexico	58 (2.1)	37 (2.2)	5 (0.7)
New York	65 (2.6)	32 (2.4)	3 (1.1)
North Carolina	63 (2.6)	34 (2.4)	3 (0.7)
North Dakota	68 (2.2)	31 (2.3)	1 (0.2)
Ohio	66 (2.1)	32 (2.0)	2 (0.6)
Oklahoma	65 (2.0)	35 (2.1)	1 (0.4)
Pennsylvania	67 (2.3)	30 (2.2)	3 (0.8)
Rhode Island	63 (1.9)	35 (2.0)	2 (0.5)
South Carolina	65 (1.9)	33 (1.8)	2 (0.7)
Tennessee	60 (2.4)	38 (2.5)	3 (0.7)
Texas	71 (1.9)	27 (1.7)	3 (0.7)
Utah	71 (2.0)	28 (1.9)	1 (0.4)
Virginia	65 (2.0)	33 (2.0)	2 (0.5)
West Virginia	59 (2.1)	39 (2.2)	2 (0.6)
Wisconsin	67 (3.3)	31 (2.7)	2 (0.9)
Wyoming	69 (1.7)	29 (1.7)	2 (0.5)
TERRITORIES			
Guam	49 (3.0)	39 (3.4)	12 (2.0)
Virgin Islands	38 (2.7)	46 (2.9)	16 (1.7)

The regular constructed-response question shown below is from the content dimension of Data Analysis, Probability, and Statistics. Students were given one outcome for the experiment of selecting two marbles from a bag of yellow and blue marbles and asked to generate the remaining outcomes. To receive credit for this question, students had to list all three remaining possible outcomes. If they listed one or more of the outcomes more than once, that was accepted as correct, providing they had listed among their responses the three unlisted outcomes (y,y), (b,y), and (b,b). Nationally, 24 percent of the grade 4 students and 59 percent of the grade 8 students correctly answered this item.

EXAMPLE 3: Data Analysis, Statistics, and Probability

Steve was asked to pick two marbles from a bag of yellow marbles and blue marbles. One possible result was one yellow marble first and one blue marble second. He wrote this result in the table below. List all of the other possible results that Steve could get.

Overall Percent Correct*
Grade 4 -- 24 (1.5)
Grade 8 -- 59 (1.3)

y stands for one yellow marble.

b stands for one blue marble.

First Marble	Second Marble
y	b
y	y
b	b
b	y

* The standard errors of the estimated percentages appear in parentheses.

The national results in TABLE 1.5 indicate considerable improvement between grade 4 and grade 8 for each of the various subpopulations. The state-by-state results in TABLE 1.6 show performance for grade 4 students ranged from 6 to 37 percent correct. At grade 8, the percentages of correct responses ranged from 22 to 75 percent correct. Compared with the results for questions in other content areas, success on questions such as this data analysis item may be dependent on curricular coverage of the topic. Of the five mathematics content areas assessed by NAEP, teachers of fourth and eighth graders reported the least instructional emphasis on the area of data analysis, probability, and statistics.[10] However, NAEP results reflect that students are both capable of understanding and dealing with this content area.

[10] *Data Compendium for the NAEP 1992 Mathematics Assessment of the Nation and the States* (Washington, DC: National Center for Education Statistics, U.S. Government Printing Office, 1993).

TABLE 1.5 National Results for Demographic Subgroups for the Regular Constructed-Response Task, "Sampling the Yellow and Blue Marbles"

Grade 4

	Correct	Incorrect	No Response
Nation	24 (1.5)	64 (1.6)	12 (1.1)
Northeast	31 (3.3)	60 (2.8)	9 (2.1)
Southeast	17 (2.3)	68 (3.4)	14 (2.6)
Central	30 (2.7)	59 (2.8)	12 (1.1)
West	19 (3.4)	68 (3.4)	13 (2.0)
White	30 (1.7)	60 (1.7)	10 (1.0)
Black	4 (1.8)	78 (3.8)	18 (3.6)
Hispanic	9 (2.7)	72 (3.7)	19 (3.3)
Male	25 (1.8)	62 (1.9)	12 (1.2)
Female	22 (2.0)	66 (2.0)	12 (1.4)
Advantaged Urban	44 (4.7)	49 (3.7)	7 (2.6)
Disadvantaged Urban	11 (3.0)	65 (4.6)	24 (4.8)
Extreme Rural	21 (3.7)	65 (4.7)	14 (3.0)
Other	22 (1.6)	66 (1.7)	11 (1.1)
Public	24 (1.6)	64 (1.7)	12 (1.1)
Catholic and Other Private	23 (2.9)	64 (3.4)	13 (2.1)

Grade 8

	Correct	Incorrect	No Response
Nation	59 (1.3)	30 (1.1)	11 (0.7)
Northeast	59 (3.2)	30 (2.9)	11 (1.4)
Southeast	53 (2.7)	34 (2.9)	13 (1.0)
Central	66 (2.6)	26 (2.1)	8 (1.7)
West	58 (2.1)	29 (1.3)	13 (1.6)
White	68 (1.3)	23 (1.2)	9 (1.0)
Black	33 (4.0)	49 (4.1)	18 (2.0)
Hispanic	38 (2.7)	42 (3.5)	20 (2.8)
Male	59 (1.6)	27 (1.5)	14 (1.1)
Female	59 (1.9)	32 (1.9)	9 (1.1)
Advantaged Urban	75 (3.1)	22 (2.9)	3 (1.4)
Disadvantaged Urban	34 (3.7)	44 (4.1)	22 (3.0)
Extreme Rural	65 (4.2)	30 (3.6)	4 (1.5)
Other	59 (1.4)	29 (1.3)	12 (0.9)
Public	58 (1.4)	30 (1.3)	12 (0.8)
Catholic and Other Private	67 (2.8)	26 (2.6)	8 (1.9)

The standard errors of the estimated percentages appear in parentheses. It can be said with about 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details). Percentages may not total 100 percent due to rounding error.

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment

TABLE 1.6 | **Percentage of Correct Responses to Regular Constructed-Response Question, "Sampling the Yellow and Blue Marbles"**

PUBLIC SCHOOLS	Grade 4 - 1992		
	Correct	Incorrect	No Response
NATION	24 (1.6)	64 (1.7)	12 (1.1)
Northeast	32 (3.7)	60 (3.4)	8 (2.2)
Southeast	17 (2.6)	68 (3.4)	15 (2.5)
Central	31 (2.9)	57 (3.3)	12 (1.5)
West	18 (3.5)	69 (3.5)	13 (2.2)
STATES			
Alabama	17 (1.6)	72 (2.0)	11 (1.4)
Arizona	19 (1.7)	67 (1.8)	13 (1.3)
Arkansas	17 (1.8)	69 (1.8)	14 (1.4)
California	18 (2.0)	70 (2.4)	12 (1.7)
Colorado	31 (2.5)	58 (2.3)	11 (1.3)
Connecticut	34 (2.3)	57 (2.5)	9 (1.5)
Delaware	18 (2.0)	68 (2.5)	14 (1.3)
Dist. Columbia	9 (1.4)	70 (2.2)	21 (1.8)
Florida	18 (1.8)	69 (1.9)	13 (1.2)
Georgia	23 (1.9)	63 (2.3)	14 (1.7)
Hawaii	19 (1.5)	67 (2.3)	14 (1.6)
Idaho	24 (2.0)	64 (2.1)	11 (1.3)
Indiana	27 (2.3)	64 (2.2)	9 (1.4)
Iowa	32 (1.8)	60 (2.0)	9 (1.0)
Kentucky	18 (1.6)	70 (1.8)	11 (1.5)
Louisiana	12 (1.7)	77 (2.0)	12 (1.4)
Maine	37 (2.2)	54 (2.3)	9 (1.5)
Maryland	30 (1.7)	61 (1.9)	10 (1.1)
Massachusetts	32 (2.6)	57 (2.6)	11 (1.5)
Michigan	21 (2.1)	69 (2.1)	10 (1.4)
Minnesota	31 (2.2)	59 (2.2)	10 (1.6)
Mississippi	10 (1.4)	73 (1.8)	17 (1.6)
Missouri	27 (2.1)	62 (2.5)	10 (1.6)
Nebraska	28 (2.3)	60 (2.6)	12 (2.0)
New Hampshire	35 (2.3)	54 (2.5)	10 (1.5)
New Jersey	30 (2.4)	59 (2.4)	11 (1.3)
New Mexico	19 (1.8)	67 (2.1)	14 (1.8)
New York	25 (2.4)	64 (2.7)	11 (1.8)
North Carolina	18 (1.3)	71 (1.3)	11 (1.3)
North Dakota	32 (1.8)	60 (2.1)	8 (1.3)
Ohio	28 (2.2)	63 (2.3)	9 (1.2)
Oklahoma	23 (1.9)	68 (2.0)	9 (1.3)
Pennsylvania	27 (2.0)	62 (1.9)	11 (1.3)
Rhode Island	21 (2.1)	62 (2.3)	16 (1.6)
South Carolina	16 (1.5)	70 (2.1)	14 (1.8)
Tennessee	18 (1.9)	66 (2.0)	15 (1.5)
Texas	18 (1.7)	68 (2.0)	15 (1.6)
Utah	21 (1.9)	67 (2.1)	12 (1.3)
Virginia	25 (1.9)	65 (1.8)	9 (1.2)
West Virginia	16 (1.6)	71 (2.1)	14 (1.4)
Wisconsin	34 (2.6)	58 (2.6)	8 (1.2)
Wyoming	29 (1.9)	58 (2.2)	13 (1.4)
TERRITORY			
Guam	6 (1.0)	84 (2.0)	11 (1.7)

The standard errors of the estimated percentages appear in parentheses. It can be said with 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details). Percentages may not total 100 percent due to rounding error.

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment.

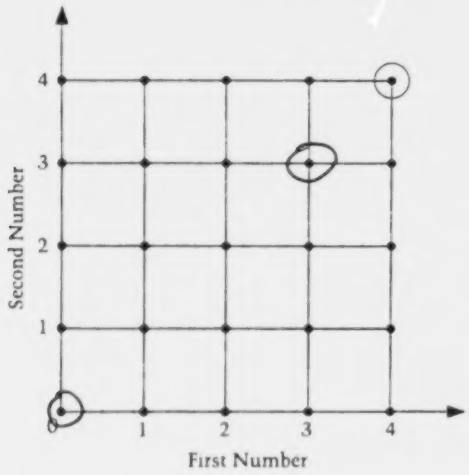
TABLE 1.6 | **Percentage of Correct Responses to Regular Constructed-Response Question, "Sampling the Yellow and Blue Marbles" (continued)**

PUBLIC SCHOOLS	Grade 8 - 1992		
	Correct	Incorrect	No Response
NATION	58 (1.4)	30 (1.3)	12 (0.8)
Northeast	58 (4.0)	29 (3.6)	12 (1.6)
Southeast	52 (2.9)	34 (3.2)	14 (1.2)
Central	66 (2.8)	26 (2.2)	8 (1.7)
West	57 (2.3)	30 (1.3)	13 (2.0)
STATES			
Alabama	45 (2.4)	38 (2.3)	17 (1.7)
Arizona	59 (2.3)	30 (1.9)	11 (1.5)
Arkansas	50 (2.4)	36 (2.0)	14 (1.6)
California	55 (2.3)	30 (2.3)	15 (1.9)
Colorado	68 (2.0)	25 (1.7)	8 (1.1)
Connecticut	64 (2.6)	29 (2.2)	7 (1.3)
Delaware	60 (2.8)	28 (2.3)	13 (1.8)
Dist. Columbia	33 (3.3)	43 (3.0)	24 (2.3)
Florida	56 (2.4)	31 (1.8)	13 (1.4)
Georgia	51 (2.3)	37 (2.3)	12 (1.3)
Hawaii	45 (2.0)	38 (2.3)	17 (1.9)
Idaho	65 (1.8)	28 (1.7)	7 (1.1)
Indiana	68 (1.7)	24 (1.5)	8 (1.1)
Iowa	75 (1.8)	21 (1.5)	3 (0.7)
Kentucky	56 (2.4)	33 (2.2)	11 (1.3)
Louisiana	48 (2.7)	37 (2.3)	15 (1.6)
Maine	72 (2.2)	25 (2.1)	3 (0.5)
Maryland	62 (2.2)	28 (2.1)	10 (1.5)
Massachusetts	61 (1.8)	29 (2.0)	10 (1.3)
Michigan	64 (2.2)	27 (1.7)	9 (1.3)
Minnesota	72 (2.1)	22 (1.8)	5 (1.1)
Mississippi	43 (2.7)	39 (2.4)	18 (1.7)
Missouri	64 (2.3)	28 (2.0)	8 (1.3)
Nebraska	69 (2.1)	24 (1.9)	6 (1.4)
New Hampshire	71 (2.0)	22 (1.8)	7 (1.1)
New Jersey	65 (2.4)	26 (1.8)	9 (1.1)
New Mexico	53 (1.9)	34 (1.7)	13 (1.5)
New York	60 (2.6)	30 (2.1)	10 (1.6)
North Carolina	58 (2.2)	33 (2.2)	9 (1.3)
North Dakota	71 (2.1)	26 (2.0)	3 (0.8)
Ohio	62 (1.9)	27 (1.9)	11 (1.2)
Oklahoma	62 (2.1)	30 (2.1)	8 (1.3)
Pennsylvania	63 (2.8)	29 (2.4)	8 (1.3)
Rhode Island	63 (2.2)	26 (2.0)	11 (1.3)
South Carolina	54 (2.1)	35 (1.8)	10 (1.3)
Tennessee	54 (1.8)	34 (1.7)	11 (1.3)
Texas	56 (2.1)	33 (1.9)	11 (1.3)
Utah	67 (2.0)	25 (1.8)	8 (1.0)
Virginia	60 (2.3)	31 (2.1)	9 (1.3)
West Virginia	56 (1.9)	34 (2.0)	10 (1.4)
Wisconsin	70 (2.4)	25 (2.2)	6 (1.2)
Wyoming	69 (1.8)	25 (1.7)	5 (0.9)
TERRITORIES			
Guam	29 (2.2)	40 (2.4)	31 (2.0)
Virgin Islands	22 (2.3)	34 (2.5)	44 (2.3)

The following question asked students to locate two other points on a coordinate grid having coordinates where the first and second components were equal. To receive credit for their answer, students needed to circle two or more of the points (0,0), (1,1), (2,2), or (3,3) on the grid. Nationwide, 38 percent of the grade 4 students and 75 percent of the grade 8 students correctly identified at least two of the required points. However, as shown in TABLE 1.7, performance varied considerably across subgroups for this question, which measures concepts that form underpinnings of algebra.

EXAMPLE 4: Algebra and Functions

On the grid below, the dot at (4, 4) is circled. Circle two other dots where the first number is equal to the second number.



Overall Percent Correct *
Grade 4 -- 38 (1.6)
Grade 8 -- 75 (1.3)

(One of the possible answers)

* The standard errors of the estimated percentages appear in parentheses.

At grade 4, the state-by-state percentages of correct responses ranged from 18 to 52 percent and at grade 8 from 40 to 88 percent (see TABLE 1.8).

TABLE 1.7 National Results for Demographic Subgroups for the Regular Constructed-Response Task, "Find Points on a Grid"

Grade 4

	Correct	Incorrect	No Response
Nation	38 (1.6)	51 (1.3)	10 (1.0)
Northeast	36 (2.4)	55 (3.2)	10 (1.9)
Southeast	33 (2.9)	55 (2.0)	12 (2.4)
Central	45 (4.4)	46 (3.6)	9 (2.1)
West	39 (2.5)	50 (1.6)	11 (1.3)
White	47 (2.2)	45 (1.8)	8 (1.1)
Black	13 (2.1)	70 (3.1)	17 (2.6)
Hispanic	15 (2.2)	68 (3.0)	17 (2.3)
Male	40 (2.0)	48 (1.8)	12 (1.3)
Female	38 (2.4)	54 (2.1)	8 (1.1)
Advantaged Urban	55 (2.9)	42 (3.3)	3 (1.2)
Disadvantaged Urban	18 (3.5)	62 (4.2)	20 (4.2)
Extreme Rural	35 (4.6)	52 (3.3)	14 (3.4)
Other	39 (1.9)	51 (1.8)	10 (1.1)
Public	38 (1.8)	51 (1.5)	11 (1.1)
Catholic and Other Private	42 (2.1)	51 (2.1)	7 (1.5)

Grade 8

	Correct	Incorrect	No Response
Nation	75 (1.3)	21 (1.3)	4 (0.6)
Northeast	72 (3.8)	24 (2.9)	4 (1.4)
Southeast	75 (2.9)	23 (2.7)	3 (1.0)
Central	81 (2.1)	16 (2.1)	3 (0.8)
West	73 (2.1)	22 (2.5)	6 (1.5)
White	81 (1.5)	16 (1.4)	3 (0.7)
Black	59 (3.8)	36 (3.6)	6 (1.9)
Hispanic	58 (2.9)	37 (2.9)	5 (1.5)
Male	76 (1.9)	20 (1.7)	5 (0.9)
Female	75 (1.5)	22 (1.5)	3 (0.6)
Advantaged Urban	80 (4.9)	15 (4.0)	4 (1.5)
Disadvantaged Urban	60 (3.2)	36 (2.7)	5 (1.7)
Extreme Rural	80 (5.1)	20 (5.1)	0 (0.0)
Other	76 (1.2)	20 (1.2)	4 (0.7)
Public	75 (1.4)	21 (1.4)	4 (0.6)
Catholic and Other Private	78 (2.3)	20 (2.2)	3 (0.9)

The standard errors of the estimated percentages appear in parentheses. It can be said with about 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details). Percentages may not total 100 percent due to rounding error.

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment

TABLE 1.8 | **Percentage of Correct Responses to Regular Constructed-Response Question, "Find Points on a Grid"**

PUBLIC SCHOOLS	Grade 4 - 1992		
	Correct	Incorrect	No Response
NATION	38 (1.8)	51 (1.5)	11 (1.1)
Northeast	36 (2.5)	55 (3.7)	9 (2.1)
Southeast	32 (3.2)	56 (2.0)	12 (2.7)
Central	45 (5.0)	45 (4.2)	10 (2.5)
West	38 (2.9)	50 (2.1)	12 (1.4)
STATES			
Alabama	26 (2.3)	63 (2.5)	12 (1.5)
Arizona	37 (1.8)	52 (2.0)	11 (1.3)
Arkansas	33 (1.9)	57 (2.0)	10 (1.2)
California	34 (2.1)	50 (2.3)	16 (1.4)
Colorado	45 (2.1)	48 (2.1)	7 (1.0)
Connecticut	42 (1.9)	51 (1.8)	7 (1.2)
Delaware	34 (2.5)	57 (2.1)	9 (1.7)
Dist. Columbia	16 (1.2)	68 (1.8)	16 (1.5)
Florida	38 (2.6)	53 (2.5)	10 (1.3)
Georgia	36 (2.1)	55 (1.9)	9 (1.1)
Hawaii	29 (2.1)	60 (2.2)	11 (1.5)
Idaho	37 (2.6)	54 (2.3)	9 (1.2)
Indiana	37 (2.1)	56 (2.3)	7 (1.3)
Iowa	46 (2.0)	47 (1.9)	7 (1.1)
Kentucky	30 (2.3)	61 (2.3)	9 (1.3)
Louisiana	24 (2.0)	64 (2.2)	12 (1.3)
Maine	46 (2.5)	49 (2.4)	5 (1.0)
Maryland	38 (1.7)	55 (1.9)	8 (1.2)
Massachusetts	39 (2.3)	52 (2.6)	9 (1.3)
Michigan	33 (2.3)	60 (2.5)	8 (1.1)
Minnesota	45 (2.1)	46 (2.2)	9 (1.1)
Mississippi	21 (1.5)	65 (2.0)	14 (1.7)
Missouri	48 (2.2)	46 (2.2)	6 (1.1)
Nebraska	42 (2.7)	51 (2.6)	7 (1.2)
New Hampshire	47 (2.4)	45 (2.4)	8 (1.4)
New Jersey	50 (2.6)	45 (2.5)	5 (1.5)
New Mexico	31 (2.9)	59 (3.1)	11 (1.4)
New York	39 (2.3)	54 (2.5)	8 (1.3)
North Carolina	34 (2.3)	58 (2.4)	8 (1.1)
North Dakota	43 (2.2)	51 (2.2)	6 (1.0)
Ohio	36 (2.0)	56 (1.9)	8 (1.4)
Oklahoma	39 (2.4)	56 (2.4)	5 (1.1)
Pennsylvania	40 (2.1)	53 (1.9)	7 (1.2)
Rhode Island	32 (2.5)	58 (2.3)	10 (1.4)
South Carolina	31 (1.8)	60 (1.9)	9 (1.0)
Tennessee	28 (2.1)	62 (2.1)	10 (1.4)
Texas	34 (2.0)	58 (1.9)	8 (1.2)
Utah	39 (2.1)	53 (2.0)	8 (1.1)
Virginia	36 (2.4)	58 (2.2)	6 (0.9)
West Virginia	33 (2.0)	57 (2.1)	9 (1.1)
Wisconsin	52 (2.2)	43 (2.1)	5 (0.9)
Wyoming	40 (2.2)	54 (2.3)	6 (1.0)
TERRITORY			
Guam	18 (2.1)	70 (2.5)	12 (1.4)

The standard errors of the estimated percentages appear in parentheses. It can be said with 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details). Percentages may not total 100 percent due to rounding error.

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment.

BEST COPY AVAILABLE

TABLE 1.8 | **Percentage of Correct Responses to Regular Constructed-Response Question, "Find Points on a Grid" (continued)**

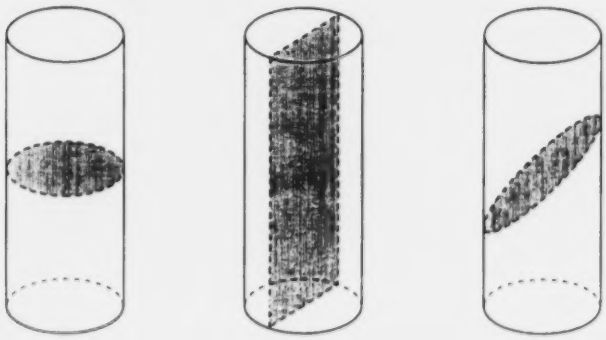
PUBLIC SCHOOLS	Grade 8 - 1992		
	Correct	Incorrect	No Response
NATION	75 (1.4)	21 (1.4)	4 (0.6)
Northeast	71 (4.7)	25 (3.5)	4 (1.7)
Southeast	75 (3.1)	23 (2.9)	2 (0.9)
Central	81 (2.3)	17 (2.3)	3 (0.9)
West	73 (2.2)	22 (2.7)	6 (1.6)
STATES			
Alabama	66 (2.1)	30 (1.9)	4 (0.9)
Arizona	74 (2.2)	22 (1.9)	4 (0.8)
Arkansas	69 (1.8)	27 (2.0)	4 (1.0)
California	71 (2.4)	24 (2.0)	6 (1.1)
Colorado	79 (1.5)	18 (1.4)	3 (0.7)
Connecticut	80 (1.7)	17 (1.5)	3 (1.0)
Delaware	77 (2.2)	20 (2.2)	3 (0.8)
Dist. Columbia	50 (2.4)	43 (2.4)	7 (1.1)
Florida	75 (2.0)	19 (2.1)	6 (1.0)
Georgia	71 (2.1)	26 (1.9)	3 (0.6)
Hawaii	65 (2.3)	29 (2.3)	5 (0.9)
Idaho	85 (1.4)	13 (1.4)	2 (0.6)
Indiana	81 (1.7)	16 (1.5)	2 (0.8)
Iowa	88 (1.4)	12 (1.4)	0 (0.3)
Kentucky	74 (1.6)	25 (1.5)	2 (0.6)
Louisiana	65 (2.2)	30 (2.0)	5 (1.0)
Maine	85 (1.6)	14 (1.5)	1 (0.4)
Maryland	72 (2.0)	24 (1.8)	4 (0.8)
Massachusetts	82 (1.6)	16 (1.5)	2 (0.6)
Michigan	77 (2.1)	19 (1.8)	4 (0.8)
Minnesota	84 (1.2)	14 (1.3)	2 (0.6)
Mississippi	60 (2.3)	36 (2.2)	4 (0.6)
Missouri	83 (1.7)	14 (1.6)	3 (0.6)
Nebraska	88 (1.7)	10 (1.4)	2 (0.6)
New Hampshire	85 (1.5)	14 (1.5)	1 (0.3)
New Jersey	79 (2.0)	20 (2.0)	1 (0.5)
New Mexico	75 (2.2)	21 (2.0)	4 (0.7)
New York	76 (2.8)	20 (2.4)	4 (1.0)
North Carolina	72 (1.8)	24 (1.6)	4 (0.7)
North Dakota	86 (1.9)	13 (1.9)	1 (0.4)
Ohio	80 (2.0)	17 (1.9)	3 (0.7)
Oklahoma	82 (1.9)	16 (1.9)	2 (0.7)
Pennsylvania	80 (2.1)	19 (1.9)	1 (0.5)
Rhode Island	80 (1.9)	16 (1.8)	4 (0.8)
South Carolina	69 (1.9)	28 (1.9)	3 (0.8)
Tennessee	72 (2.0)	25 (1.9)	3 (0.7)
Texas	77 (1.9)	19 (1.6)	3 (0.8)
Utah	86 (1.4)	13 (1.4)	2 (0.5)
Virginia	76 (1.9)	22 (1.8)	2 (0.7)
West Virginia	76 (2.1)	22 (1.9)	2 (0.5)
Wisconsin	84 (1.6)	14 (1.5)	2 (0.6)
Wyoming	82 (1.8)	16 (1.7)	2 (0.5)
TERRITORIES			
Guam	55 (2.8)	35 (2.7)	10 (1.6)
Virgin Islands	40 (2.5)	43 (3.0)	17 (2.1)

The geometry question presented below, given only to eighth-grade students, required them to identify the shape of various cross sections of a cylinder. To receive credit for their answers, students needed to identify all three cross sections: circle; rectangle, parallelogram, or quadrilateral; and oval or ellipse. This type of question reflects one aspect of the added emphasis given to spatial visualization in school mathematics. The increased emphasis on spatial awareness and its relationship to shapes and properties is a central portion of the middle-grade goals articulated in *The NCTM Standards*. Such understanding is central to the application of geometry in the solution of many real-world problems, especially those involved in engineering and several vocational or technical areas. For the nation, 48 percent of grade 8 students correctly completed the question.

EXAMPLE 5: Geometry

Each of the cylinders shown below was cut in a different way. The shaded part shows the shape of the cut. Under each figure, write the name of the shape of the cut.

Overall Percent Correct *
Grade 8 -- 48 (1.3)



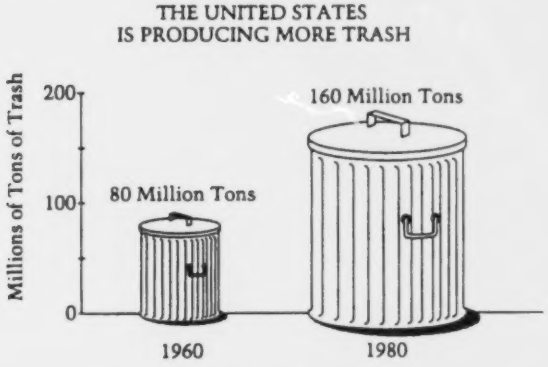
Answer: circle Answer: rectangle Answer: oval

*The standard errors of the estimated percentages appear in parentheses.

The "Trash Cans" question, which was in the Data Analysis, Probability, and Statistics content area, required eighth-grade students to examine a misleading pictograph and explain why the data display was misleading. To receive credit for a correct response, students needed to note that the 1980 can would hold more than twice the 1960 can or that both the width and height of the can had been doubled. (In particular, doubling the dimensions of the can would lead to an eightfold increase in the volume of the can, because doubling the radius [or diameter] results in a fourfold increase when the radius is squared in $v=\pi r^2 h$.) However, even though the general rather than the specific answer was scored correct, student performance at the national level was quite low, with 8 percent of the eighth-grade students providing an acceptable response.

The ability to read data from a graph, noting the correctness of the graph and the implied comparisons, is an important consumer skill. The ability to detect errors of the type presented in this question is an important outcome of the data analysis/quantitative literacy aspect of the school mathematics curriculum. While some students seem to have developed this critical skill, the results indicate that the vast majority have little conception of the effects that such visual representations can have on the possible interpretations of the data.

EXAMPLE 6: Data Analysis, Statistics, and Probability



The pictograph shown above is misleading. Explain why.

Answer: Both the width and the height of the 1980 can have been doubled. Only the height should have been doubled.

Overall Percent Correct *
Grade 8 -- 8 (0.8)

(One of the possible answers)

*The standard errors of the estimated percentages appear in parentheses.

The subgroup results for eighth graders on both the "Shapes Cut from Cylinders" question and the "Misleading Pictograph of Trash Cans" question are found in TABLE 1.9. Although both questions were difficult for students, especially the pictograph task, differences among subgroups were somewhat less pronounced than for some of the previous example questions. The state results for the two questions in TABLE 1.10 reveal that the percentages of correct responses ranged from 36 to 60 percent for "Cylinders" and from 1 to 16 percent for the "Trash Cans" question.

TABLE 1.9 National Results for Demographic Subgroups for the Regular Constructed-Response Tasks, "Shapes Cut from Cylinders" and "Misleading Pictograph of Trash Cans"

Grade 8

	Shapes Cut from Cylinders			Misleading Pictograph of Trash Cans		
	Correct	Incorrect	No Response	Correct	Incorrect	No Response
Nation	48 (1.3)	45 (1.2)	7 (0.8)	8 (0.8)	86 (1.1)	6 (0.9)
Northeast	46 (2.7)	48 (2.1)	6 (1.1)	14 (2.4)	78 (3.0)	8 (1.9)
Southeast	46 (2.9)	46 (2.6)	8 (2.2)	6 (0.9)	90 (1.0)	5 (0.9)
Central	53 (2.5)	42 (2.5)	5 (0.9)	10 (1.8)	87 (2.4)	3 (1.3)
West	48 (1.9)	44 (2.3)	8 (1.6)	6 (1.5)	87 (2.2)	7 (2.3)
White	52 (1.6)	42 (1.7)	5 (0.9)	10 (1.0)	84 (1.4)	6 (1.0)
Black	41 (4.0)	48 (3.1)	11 (2.7)	4 (1.5)	90 (2.5)	6 (2.2)
Hispanic	38 (3.3)	51 (3.3)	10 (2.2)	4 (1.5)	92 (1.9)	4 (1.0)
Male	44 (2.1)	48 (2.1)	8 (1.1)	10 (1.3)	84 (1.3)	6 (0.9)
Female	54 (1.5)	42 (1.4)	5 (0.9)	7 (0.9)	87 (1.7)	6 (1.3)
Advantaged Urban	54 (2.9)	42 (3.4)	3 (1.5)	12 (2.3)	84 (2.8)	4 (1.3)
Disadvantaged Urban	35 (4.3)	51 (4.3)	14 (2.8)	2 (0.9)	89 (2.1)	10 (2.2)
Extreme Rural	50 (5.2)	45 (5.0)	5 (3.0)	5 (2.4)	91 (2.3)	4 (1.7)
Other	49 (1.6)	44 (1.4)	7 (0.9)	9 (1.0)	85 (1.4)	6 (1.1)
Public	48 (1.5)	45 (1.4)	7 (0.9)	8 (0.9)	86 (1.2)	6 (0.9)
Catholic and Other Private	50 (2.5)	46 (2.5)	4 (1.0)	12 (1.7)	83 (2.2)	5 (1.4)

The standard errors of the estimated percentages appear in parentheses. It can be said with about 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details). Percentages may not total 100 percent due to rounding error.

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment

TABLE 1.10 | **Percentages of Correct Responses to Regular Constructed-Response Questions, "Shapes Cut from Cylinders" and "Misleading Pictograph of Trash Cans"**

PUBLIC SCHOOLS	Grade 8 - 1992					
	Shapes Cut from Cylinders			Misleading Pictographs of Trash Cans		
	Correct	Incorrect	No Response	Correct	Incorrect	No Response
NATION	48 (1.5)	45 (1.4)	7 (0.9)	8 (0.9)	86 (1.2)	6 (0.9)
Northeast	45 (3.5)	48 (2.9)	7 (1.4)	14 (2.7)	78 (3.0)	8 (2.0)
Southeast	46 (3.2)	46 (2.8)	8 (2.3)	5 (0.9)	90 (1.1)	5 (1.0)
Central	53 (3.1)	41 (2.9)	5 (1.1)	9 (2.2)	88 (2.7)	3 (1.4)
West	48 (1.9)	44 (2.5)	8 (1.9)	5 (1.6)	87 (2.3)	8 (2.5)
STATES						
Alabama	36 (1.8)	57 (1.7)	7 (1.1)	7 (1.4)	91 (1.3)	2 (0.6)
Arizona	51 (2.3)	43 (2.2)	6 (1.0)	7 (1.2)	87 (1.4)	6 (0.9)
Arkansas	42 (2.0)	52 (2.0)	7 (1.2)	6 (1.1)	87 (1.3)	7 (1.0)
California	44 (2.3)	47 (2.3)	9 (1.3)	7 (1.2)	88 (1.5)	6 (1.3)
Colorado	50 (2.3)	43 (2.2)	7 (1.0)	9 (1.1)	86 (1.2)	5 (1.1)
Connecticut	52 (2.0)	44 (2.2)	4 (0.7)	10 (1.4)	86 (1.4)	4 (0.9)
Delaware	50 (2.8)	40 (2.7)	10 (1.4)	10 (1.8)	84 (2.0)	6 (1.3)
Dist. Columbia	41 (2.4)	48 (2.6)	11 (1.7)	5 (1.6)	86 (2.2)	8 (1.6)
Florida	45 (2.5)	45 (2.5)	10 (1.3)	5 (1.1)	88 (1.5)	7 (1.3)
Georgia	45 (2.4)	47 (2.4)	7 (0.9)	6 (1.3)	90 (1.5)	3 (0.7)
Hawaii	48 (1.9)	41 (2.1)	11 (1.4)	6 (1.2)	87 (1.7)	7 (1.2)
Idaho	50 (2.1)	44 (2.0)	6 (1.0)	11 (1.4)	85 (1.5)	4 (0.8)
Indiana	55 (2.2)	40 (2.1)	5 (0.9)	8 (1.2)	89 (1.4)	4 (1.0)
Iowa	54 (2.3)	43 (2.1)	3 (0.7)	13 (1.5)	83 (1.5)	3 (0.6)
Kentucky	45 (2.2)	49 (2.4)	6 (1.1)	8 (1.4)	88 (1.6)	4 (1.0)
Louisiana	40 (2.5)	51 (2.3)	9 (1.3)	4 (0.9)	91 (1.3)	5 (1.1)
Maine	56 (2.0)	38 (2.2)	5 (1.1)	10 (1.5)	86 (1.8)	3 (0.9)
Maryland	48 (2.8)	46 (2.6)	6 (0.9)	11 (1.5)	86 (1.6)	3 (0.9)
Massachusetts	56 (2.2)	38 (2.3)	6 (1.1)	16 (2.3)	80 (2.4)	4 (0.9)
Michigan	52 (2.0)	43 (1.8)	5 (1.1)	11 (1.5)	86 (1.4)	4 (0.9)
Minnesota	52 (2.6)	44 (2.3)	5 (0.9)	11 (1.4)	84 (1.7)	5 (0.9)
Mississippi	34 (2.2)	56 (1.9)	10 (1.5)	4 (0.9)	92 (1.0)	3 (1.0)
Missouri	56 (2.2)	38 (2.0)	6 (1.0)	12 (1.6)	83 (1.8)	6 (1.0)
Nebraska	50 (2.6)	44 (2.5)	6 (1.3)	15 (1.5)	79 (1.6)	6 (1.0)
New Hampshire	60 (2.3)	35 (2.2)	5 (0.9)	16 (2.0)	80 (2.1)	4 (0.9)
New Jersey	51 (2.6)	45 (2.9)	5 (1.0)	11 (1.4)	87 (1.6)	2 (0.7)
New Mexico	47 (2.1)	46 (2.2)	7 (1.0)	4 (1.0)	91 (1.3)	5 (1.0)
New York	54 (2.6)	38 (2.5)	8 (1.2)	9 (1.2)	86 (1.4)	5 (1.1)
North Carolina	45 (2.1)	49 (2.1)	6 (1.0)	8 (1.2)	89 (1.4)	4 (0.8)
North Dakota	50 (2.6)	47 (2.6)	3 (0.8)	16 (1.8)	81 (1.8)	3 (0.9)
Ohio	54 (2.4)	41 (2.2)	5 (1.0)	11 (1.4)	85 (1.6)	4 (0.8)
Oklahoma	49 (2.6)	45 (2.5)	6 (1.1)	10 (1.3)	88 (1.4)	2 (0.7)
Pennsylvania	56 (1.9)	40 (1.7)	4 (0.9)	13 (1.4)	84 (1.5)	3 (0.6)
Rhode Island	52 (3.1)	42 (3.0)	6 (1.1)	11 (1.7)	85 (1.9)	4 (1.0)
South Carolina	46 (2.0)	48 (2.1)	6 (0.9)	6 (1.0)	89 (1.4)	4 (1.0)
Tennessee	48 (1.8)	46 (1.8)	6 (0.9)	8 (1.2)	89 (1.4)	3 (0.8)
Texas	45 (2.6)	48 (2.5)	7 (1.3)	8 (1.2)	87 (1.3)	6 (0.9)
Utah	53 (2.3)	41 (2.2)	6 (1.2)	13 (1.5)	81 (2.0)	5 (1.1)
Virginia	53 (2.3)	42 (2.3)	5 (0.8)	9 (1.4)	87 (1.7)	4 (0.9)
West Virginia	46 (2.3)	47 (2.3)	8 (1.0)	8 (1.4)	87 (1.6)	5 (1.1)
Wisconsin	51 (2.3)	44 (2.6)	5 (1.0)	11 (1.5)	86 (1.5)	4 (0.6)
Wyoming	51 (2.0)	44 (1.8)	5 (1.0)	8 (1.5)	88 (1.6)	3 (0.9)
TERRITORIES						
Guam	41 (2.6)	45 (2.9)	14 (1.8)	1 (0.7)	95 (1.6)	4 (1.5)
Virgin Islands	40 (3.1)	45 (3.0)	15 (2.1)	1 (0.9)	89 (2.1)	10 (1.9)

The standard errors of the estimated percentages appear in parentheses. It can be said with 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details). Percentages may not total 100 percent due to rounding error.

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment.

The following algebra and functions question also was administered only to grade 8 students. It asked students to write two numbers that could be put in the □ to make the number sentence 54 < 3 x □ true. Grade 8 students who either gave two numbers greater than 18 or communicated a generalization indicating that any number greater than 18 would suffice were given credit for a correct response. (Students giving only one number greater than 18 were not given credit.) Nationally, approximately half of the eighth graders provided an acceptable response. However, the results presented in TABLE 1.11 indicate large performance gaps for Black and Hispanic students compared to White students on this task, and between advantaged urban and disadvantaged urban students. At the state level, performance of grade 8 students ranged from 22 to 64 percent correct, as shown in TABLE 1.20.

EXAMPLE 7: Algebra and Functions

$54 < 3 \times \square$

Write two numbers that could be put in the □ to make the number sentence above true.

Answer: 19, 20 (One of the possible answers)

Overall Percent Correct*
Grade 8 -- 49 (1.6)

*The standard errors of the estimated percentages appear in parentheses.

TABLE 1.11 National Results for Demographic Subgroups for the Regular Constructed-Response Task, "Two Numbers Make Number Sentence True"

Grade 8

	Correct	Incorrect	No Response
Nation	49 (1.6)	41 (1.4)	10 (0.8)
Northeast	48 (5.4)	44 (4.9)	8 (1.1)
Southeast	43 (3.1)	46 (2.6)	11 (2.0)
Central	56 (2.9)	36 (2.3)	9 (1.7)
West	50 (2.3)	38 (1.8)	12 (1.3)
White	57 (1.9)	36 (1.8)	8 (0.8)
Black	30 (3.6)	56 (3.1)	13 (2.8)
Hispanic	25 (3.2)	53 (3.3)	22 (2.7)
Male	49 (2.0)	41 (1.8)	11 (1.2)
Female	50 (2.0)	41 (1.8)	9 (1.0)
Advantaged Urban	63 (5.4)	33 (4.7)	4 (1.6)
Disadvantaged Urban	26 (4.3)	55 (4.4)	19 (2.4)
Extreme Rural	55 (4.5)	37 (4.0)	8 (2.0)
Other	49 (1.6)	41 (1.4)	10 (0.9)
Public	48 (1.7)	42 (1.6)	10 (0.9)
Catholic and Other Private	56 (3.2)	36 (2.6)	8 (1.8)

The standard errors of the estimated percentages appear in parentheses. It can be said with about 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details). Percentages may not total 100 percent due to rounding error.

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment

TABLE 1.12 | **Percentage of Correct Responses to Regular Constructed-Response Question, "Two Numbers Make Number Sentence True"**

PUBLIC SCHOOLS	Grade 8 - 1992		
	Correct	Incorrect	No Response
NATION	48 (1.7)	42 (1.6)	10 (0.9)
Northeast	46 (6.1)	46 (5.6)	8 (1.6)
Southeast	42 (3.1)	47 (2.6)	11 (2.2)
Central	55 (3.2)	36 (2.8)	9 (1.9)
West	50 (2.3)	38 (1.9)	12 (1.3)
STATES			
Alabama	40 (2.3)	47 (2.1)	13 (1.8)
Arizona	47 (1.8)	42 (1.7)	11 (1.3)
Arkansas	42 (2.2)	47 (2.3)	10 (1.4)
California	45 (2.3)	41 (2.2)	14 (1.5)
Colorado	54 (2.4)	39 (2.2)	7 (1.0)
Connecticut	55 (2.2)	35 (2.0)	10 (1.4)
Delaware	46 (2.3)	42 (2.5)	11 (1.7)
Dist. Columbia	32 (2.8)	45 (2.5)	22 (2.0)
Florida	48 (2.2)	39 (2.0)	12 (1.4)
Georgia	49 (2.3)	42 (2.3)	9 (1.2)
Hawaii	38 (2.0)	45 (2.1)	17 (1.7)
Idaho	55 (2.1)	37 (1.8)	8 (1.0)
Indiana	46 (1.9)	45 (2.4)	9 (1.5)
Iowa	64 (2.0)	30 (1.8)	6 (0.9)
Kentucky	44 (2.2)	46 (1.9)	10 (1.2)
Louisiana	41 (2.2)	48 (2.1)	11 (1.5)
Maine	56 (2.3)	36 (2.0)	9 (1.1)
Maryland	45 (2.0)	43 (2.1)	12 (1.5)
Massachusetts	54 (2.1)	37 (1.9)	9 (1.2)
Michigan	52 (2.4)	38 (2.3)	11 (1.4)
Minnesota	63 (2.4)	31 (2.0)	6 (1.1)
Mississippi	35 (2.4)	55 (2.3)	10 (1.5)
Missouri	57 (2.1)	38 (2.3)	5 (1.1)
Nebraska	57 (2.6)	37 (2.4)	6 (0.8)
New Hampshire	55 (2.3)	38 (2.3)	7 (1.0)
New Jersey	55 (2.5)	38 (2.4)	6 (1.0)
New Mexico	37 (2.2)	51 (1.9)	11 (1.2)
New York	52 (2.8)	38 (3.0)	10 (2.0)
North Carolina	43 (2.1)	49 (2.2)	9 (1.2)
North Dakota	61 (2.4)	36 (2.6)	3 (0.8)
Ohio	50 (2.5)	42 (2.2)	8 (1.4)
Oklahoma	54 (2.4)	36 (2.2)	11 (1.4)
Pennsylvania	56 (2.5)	37 (2.0)	7 (1.2)
Rhode Island	50 (2.5)	42 (2.7)	8 (1.3)
South Carolina	46 (2.0)	45 (2.3)	9 (1.3)
Tennessee	40 (2.0)	48 (2.0)	11 (1.3)
Texas	49 (2.4)	40 (2.2)	10 (1.1)
Utah	55 (1.9)	37 (1.9)	8 (1.2)
Virginia	51 (2.2)	40 (1.9)	9 (1.3)
West Virginia	40 (2.1)	49 (2.0)	11 (1.2)
Wisconsin	55 (2.5)	37 (1.8)	8 (1.4)
Wyoming	53 (1.9)	41 (1.9)	7 (1.0)
TERRITORIES			
Guam	28 (2.1)	50 (2.6)	22 (2.3)
Virgin Islands	22 (2.2)	37 (2.6)	41 (3.0)

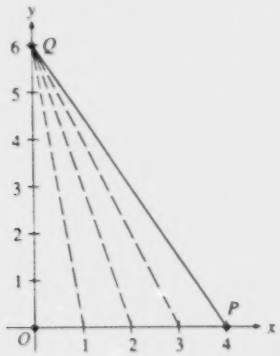
The standard errors of the estimated percentages appear in parentheses. It can be said with 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details). Percentages may not total 100 percent due to rounding error.

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment.

BEST COPY AVAILABLE

The final regular constructed-response question presented in this section was classified under the Measurement content area. This question, given to twelfth graders, asked students to consider the effect of a moving vertex on the area of a right triangle. The analysis of this situation foreshadows the nature of the use of limits in similar situations elsewhere in mathematics. As shown below and in TABLE 1.13, 29 percent of the twelfth-grade students received full credit for their responses. Fifty-one percent of the students attending schools in advantaged urban communities answered this question correctly compared to 15 percent of the students attending schools in disadvantaged urban communities. (As this item was only given at the twelfth grade, state-level data are not available.) Black and Hispanic students performed considerably below White students, and public-school students performed significantly less well than did Catholic- and other private-school students.

EXAMPLE 8: Measurement



In the figure above, point Q is fixed and point P starts at 4 and moves left along the x-axis. As P moves left along the x-axis toward O, the area of $\triangle POQ$ changes.

Use the information given to complete the table below to show how the area of $\triangle POQ$ changes as P goes from the position shown to the origin O.

x - coordinate of P	Area of $\triangle POQ$
4	12
3	9
2	6
1	3
0	0

Overall Percent Correct *
Grade 12 – 29 (1.6)

*The standard errors of the estimated percentages appear in parentheses.

TABLE 1.13 National Results for Demographic Subgroups for the Regular Constructed-Response Task, "Changing Area of Triangle"

Grade 12

	Correct	Incorrect	No Response
Nation	29 (1.6)	37 (1.7)	34 (1.5)
Northeast	38 (2.0)	32 (2.3)	31 (2.4)
Southeast	20 (2.2)	43 (2.9)	36 (3.6)
Central	29 (3.8)	35 (3.1)	36 (3.0)
West	29 (3.6)	38 (3.9)	33 (3.1)
White	33 (1.9)	35 (1.8)	32 (1.8)
Black	12 (2.6)	49 (3.0)	39 (2.9)
Hispanic	16 (3.9)	34 (4.3)	49 (5.5)
Male	30 (2.0)	38 (2.3)	32 (1.8)
Female	28 (2.1)	36 (2.3)	36 (2.4)
Advantaged Urban	51 (5.8)	30 (4.1)	19 (3.8)
Disadvantaged Urban	15 (2.4)	42 (3.7)	43 (4.1)
Extreme Rural	19 (3.8)	37 (3.8)	44 (5.5)
Other	28 (2.2)	38 (2.2)	34 (1.9)
Public	27 (1.8)	37 (1.9)	36 (1.7)
Catholic and Other Private	42 (3.1)	37 (3.2)	21 (2.6)

The standard errors of the estimated percentages appear in parentheses. It can be said with about 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details). Percentages may total 100 percent due to rounding error.

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment

Example Constructed-Response Questions Involving Rulers, Protractors, or Manipulative Materials

Some questions requiring students to construct their own responses involved use of mathematical tools -- a ruler, a protractor, or manipulative geometric shapes. The questions involving the ruler or protractor required students to make measurements or to provide a fairly accurate rendering of a geometric object. The questions involving the geometrically-shaped pieces required students to assemble them to create other geometric shapes having given properties or areas. However, none of the questions involving the geometric shapes were among the items released at the conclusion of the 1992 NAEP assessment. As a result, the following five examples of constructed-response tasks with added materials all involve a ruler or protractor.

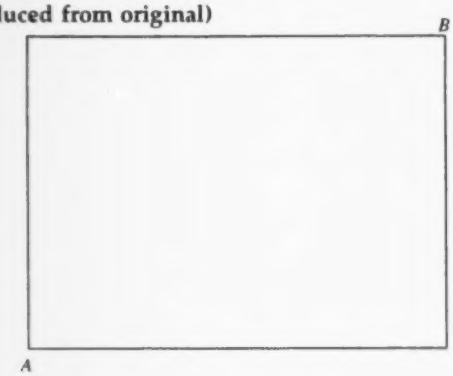
The following example question was classified in the Measurement content area. Students were shown a rectangle and first asked to measure the length of one of the longer sides to the nearest centimeter. To receive credit for their

answer, students needed to respond with a measure of 8cm. Nationally, 52 percent of the fourth graders and 71 percent of the eighth graders were able to respond correctly to the task. Students missing this question often responded with the measure of the shorter side of the rectangle or incorrectly measured the length of the longer side.

The second part of this measurement question asked students to find the length of the diagonal AB of the rectangle. Performance on this second question was about 20 percentage points higher. Among students at grade 4, 60 percent answered the item correctly, while 79 percent of the grade 8 students did. The increase in performance may be explained by the task set for students. In the first portion of the item, students were asked to measure the longer side of the rectangle, but no side was actually prescribed. In the second portion, dealing with the diagonal, students were told to measure the length of the diagonal AB. This specificity may be responsible for the increased level of performance, as some students measured the shorter side in the first portion of the question.

EXAMPLE 9: Measurement

(size reduced from original)



Overall Percent Correct*

Grade 4

Part One: 52 (1.5)
Part Two: 60 (1.2)

Grade 8

Part One: 71 (1.5)
Part Two: 79 (1.1)

Use your centimeter ruler to make the following measurements to the nearest centimeter.

What is the length in centimeters of one of the longer sides of the rectangle?

Answer: 8 centimeters

What is the length in centimeters of the diagonal from *A* to *B* ?

Answer: 10 centimeters

***The standard errors of the estimated percentages appear in parentheses.**

TABLES 1.14 and 1.15 illustrate the variation in performance that existed among demographic subpopulations of students for the nation and across the states participating in NAEP's Trial State Assessment Program. Black students and those attending schools in disadvantaged urban areas had particular difficulty in using a ruler to measure centimeters. For the states, success in measuring the longer side of the rectangle ranged from 27 to 64 percent at grade 4 and from 38 to 85 percent at grade 8. Success in measuring the diagonal ranged from 34 to 70 percent at grade 4 and from 53 to 90 percent at grade 8.

TABLE 1.14 National Results for Demographic Subgroups for the Regular Constructed Response Task, "Using Ruler to Measure Centimeters"

Grade 4

	Measure Longer Side		
	Correct	Incorrect	No Response
Nation	52 (1.5)	43 (1.4)	6 (0.7)
Northeast	52 (3.2)	42 (2.6)	7 (2.3)
Southeast	47 (2.2)	49 (2.1)	4 (1.2)
Central	56 (4.5)	40 (4.2)	3 (1.0)
West	52 (1.5)	39 (1.5)	9 (1.4)
White	57 (2.0)	38 (1.9)	5 (0.9)
Black	29 (3.5)	63 (3.6)	8 (1.7)
Hispanic	45 (3.4)	47 (3.0)	8 (1.9)
Male	51 (1.9)	42 (1.7)	7 (1.0)
Female	52 (2.5)	43 (2.5)	5 (0.9)
Advantaged Urban	66 (4.8)	31 (4.7)	2 (1.2)
Disadvantaged Urban	25 (4.0)	60 (4.1)	14 (2.4)
Extreme Rural	53 (7.5)	44 (7.4)	3 (1.4)
Other	52 (1.8)	42 (1.7)	6 (1.0)
Public	50 (1.6)	44 (1.5)	6 (0.8)
Catholic and Other Private	62 (2.8)	33 (2.4)	5 (1.3)

Grade 8

	Measure Longer Side		
	Correct	Incorrect	No Response
Nation	71 (1.5)	26 (1.3)	3 (0.5)
Northeast	73 (3.3)	24 (3.0)	3 (0.9)
Southeast	63 (3.2)	35 (2.5)	2 (1.0)
Central	77 (2.7)	20 (2.1)	3 (1.3)
West	70 (3.5)	26 (3.1)	4 (0.8)
White	78 (1.6)	19 (1.4)	2 (0.5)
Black	42 (4.1)	51 (3.6)	6 (1.7)
Hispanic	57 (3.9)	38 (4.1)	5 (1.3)
Male	71 (2.0)	26 (1.9)	3 (0.6)
Female	70 (1.8)	27 (1.6)	3 (0.7)
Advantaged Urban	82 (2.8)	15 (2.2)	2 (1.4)
Disadvantaged Urban	50 (6.1)	39 (4.6)	11 (2.6)
Extreme Rural	67 (6.5)	33 (6.5)	0 (0.3)
Other	72 (2.0)	25 (1.8)	3 (0.6)
Public	69 (1.7)	27 (1.5)	3 (0.5)
Catholic and Other Private	80 (2.2)	19 (2.2)	2 (0.9)

(Table 1.14 continued on next page)

TABLE 1.14 National Results for Demographic Subgroups for the Regular Constructed-Response Task, "Using Ruler to Measure Centimeters" (continued)

Grade 4

	Measure Diagonal		
	Correct	Incorrect	No Response
Nation	60 (1.2)	33 (1.1)	7 (0.7)
Northeast	66 (3.2)	27 (2.2)	7 (2.1)
Southeast	52 (2.6)	44 (2.0)	4 (1.3)
Central	68 (3.0)	28 (3.1)	4 (1.3)
West	57 (1.9)	33 (1.4)	10 (1.5)
White	67 (1.5)	28 (1.4)	5 (0.8)
Black	35 (3.0)	54 (3.3)	11 (2.1)
Hispanic	51 (4.2)	40 (3.3)	9 (2.1)
Male	56 (2.1)	36 (1.9)	8 (1.0)
Female	65 (1.9)	30 (1.8)	5 (0.9)
Advantaged Urban	78 (3.3)	20 (3.0)	2 (1.2)
Disadvantaged Urban	27 (3.5)	56 (3.5)	17 (3.0)
Extreme Rural	61 (7.2)	38 (6.8)	2 (1.3)
Other	62 (1.6)	31 (1.4)	7 (1.0)
Public	59 (1.3)	34 (1.2)	7 (0.8)
Catholic and Other Private	68 (2.6)	28 (2.5)	4 (1.1)

Grade 8

	Measure Diagonal		
	Correct	Incorrect	No Response
Nation	79 (1.1)	18 (1.1)	3 (0.5)
Northeast	83 (2.3)	14 (2.2)	3 (0.7)
Southeast	76 (2.2)	21 (1.4)	3 (1.2)
Central	82 (2.1)	16 (2.0)	3 (1.3)
West	75 (2.3)	21 (2.6)	4 (0.8)
White	84 (1.2)	14 (1.3)	2 (0.5)
Black	60 (4.0)	33 (3.4)	7 (1.8)
Hispanic	66 (3.2)	28 (3.3)	5 (1.2)
Male	80 (1.8)	18 (1.7)	3 (0.6)
Female	78 (1.4)	19 (1.4)	4 (0.7)
Advantaged Urban	87 (3.3)	13 (3.3)	0 (0.0)
Disadvantaged Urban	60 (5.3)	28 (4.0)	11 (2.6)
Extreme Rural	79 (2.4)	19 (2.3)	2 (1.2)
Other	80 (1.4)	18 (1.4)	3 (0.6)
Public	77 (1.2)	19 (1.2)	4 (0.6)
Catholic and Other Private	90 (1.7)	9 (1.6)	1 (0.4)

The standard errors of the estimated percentages appear in parentheses. It can be said with about 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details). When the proportion of students is either 0 percent of 100 percent, the standard error is inestimable. However, percentages 99.5 percent and greater were rounded to 100 percent and percentages 0.5 percent or less were rounded to 0 percent. Percentages may not total 100 percent due to rounding error.

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment

TABLE 1.15 | **Percentages of Correct Responses to Regular Constructed-Response Questions "Using Ruler to Measure Centimeters"**

PUBLIC SCHOOLS	Grade 4 - 1992		
	Measure Longer Side		
	Correct	Incorrect	No Response
NATION	50 (1.6)	44 (1.5)	6 (0.8)
Northeast	51 (3.5)	42 (2.9)	7 (2.6)
Southeast	45 (2.4)	51 (2.5)	4 (1.3)
Central	54 (4.9)	43 (4.7)	4 (1.0)
West	51 (1.5)	40 (1.6)	9 (1.5)
STATES			
Alabama	40 (2.1)	56 (2.2)	4 (1.0)
Arizona	51 (2.2)	43 (2.1)	5 (0.8)
Arkansas	43 (2.1)	51 (2.0)	6 (0.9)
California	45 (2.1)	46 (2.1)	9 (1.1)
Colorado	56 (2.4)	39 (2.3)	4 (0.8)
Connecticut	60 (1.9)	37 (1.8)	3 (0.7)
Delaware	48 (1.9)	49 (2.0)	4 (0.9)
Dist. Columbia	25 (2.1)	63 (2.2)	12 (1.5)
Florida	47 (2.6)	49 (2.5)	5 (0.9)
Georgia	48 (2.7)	47 (2.5)	5 (1.1)
Hawaii	48 (2.1)	47 (1.9)	5 (1.0)
Idaho	57 (2.2)	39 (2.2)	4 (0.7)
Indiana	50 (2.6)	48 (2.4)	2 (0.6)
Iowa	59 (1.8)	39 (1.7)	2 (0.7)
Kentucky	42 (2.1)	53 (2.2)	5 (0.7)
Louisiana	42 (2.9)	54 (2.7)	4 (0.8)
Maine	64 (2.7)	33 (2.6)	3 (0.9)
Maryland	49 (2.1)	47 (2.1)	4 (1.0)
Massachusetts	54 (2.6)	41 (2.8)	5 (1.2)
Michigan	55 (2.3)	40 (2.1)	5 (0.9)
Minnesota	62 (2.1)	34 (2.0)	3 (0.7)
Mississippi	36 (1.8)	58 (1.8)	5 (0.7)
Missouri	56 (2.4)	41 (2.4)	3 (0.7)
Nebraska	60 (2.7)	37 (2.7)	2 (0.7)
New Hampshire	61 (2.2)	35 (2.1)	4 (0.9)
New Jersey	56 (2.8)	39 (2.7)	5 (1.0)
New Mexico	49 (3.3)	48 (3.2)	4 (0.9)
New York	50 (2.9)	44 (2.8)	5 (1.3)
North Carolina	44 (1.6)	49 (1.8)	6 (1.1)
North Dakota	61 (2.3)	37 (2.2)	2 (0.7)
Ohio	51 (2.3)	45 (2.1)	4 (0.9)
Oklahoma	49 (2.2)	49 (2.2)	2 (0.7)
Pennsylvania	53 (2.6)	44 (2.6)	3 (0.6)
Rhode Island	46 (2.5)	47 (2.3)	8 (1.4)
South Carolina	45 (2.1)	51 (2.1)	4 (0.8)
Tennessee	40 (2.1)	55 (1.9)	5 (1.0)
Texas	49 (2.4)	49 (2.4)	2 (0.6)
Utah	59 (2.1)	38 (2.1)	3 (0.7)
Virginia	50 (2.3)	47 (2.1)	3 (1.0)
West Virginia	50 (2.0)	47 (2.0)	3 (0.7)
Wisconsin	59 (2.6)	39 (2.6)	2 (0.5)
Wyoming	60 (1.8)	35 (1.8)	5 (0.9)
TERRITORY			
Guam	27 (2.1)	63 (2.2)	10 (1.8)

The standard errors of the estimated percentages appear in parentheses. It can be said with 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details). Percentages may not total 100 percent due to rounding error.

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment.

TABLE 1.15 | **Percentages of Correct Responses to Regular Constructed-Response Questions "Using Ruler to Measure Centimeters" (continued)**

PUBLIC SCHOOLS	Grade 8 - 1992		
	Measure Longer Side		
	Correct	Incorrect	No Response
NATION	69 (1.7)	27 (1.5)	3 (0.5)
Northeast	72 (3.5)	24 (3.0)	3 (0.7)
Southeast	61 (3.6)	37 (2.9)	2 (1.2)
Central	76 (3.1)	21 (2.4)	3 (1.5)
West	69 (3.8)	26 (3.4)	5 (0.8)
STATES			
Alabama	63 (2.2)	35 (2.1)	2 (0.8)
Arizona	76 (2.0)	22 (1.9)	3 (0.7)
Arkansas	68 (1.8)	31 (1.7)	1 (0.5)
California	73 (2.0)	24 (1.8)	3 (0.7)
Colorado	81 (1.5)	17 (1.4)	2 (0.5)
Connecticut	80 (1.7)	19 (1.6)	1 (0.5)
Delaware	72 (2.0)	25 (1.8)	2 (0.6)
Dist. Columbia	45 (2.9)	45 (3.0)	10 (1.4)
Florida	68 (2.5)	30 (2.3)	3 (0.7)
Georgia	65 (2.6)	33 (2.4)	2 (0.6)
Hawaii	71 (2.0)	26 (1.8)	3 (0.6)
Idaho	80 (1.9)	18 (1.7)	2 (0.6)
Indiana	76 (1.9)	23 (1.8)	1 (0.3)
Iowa	85 (1.6)	14 (1.6)	1 (0.4)
Kentucky	71 (1.9)	27 (1.7)	2 (0.5)
Louisiana	60 (2.2)	36 (2.2)	4 (0.8)
Maine	84 (1.1)	15 (1.2)	1 (0.5)
Maryland	70 (2.0)	26 (1.8)	4 (1.0)
Massachusetts	77 (1.7)	20 (1.8)	3 (0.7)
Michigan	74 (1.7)	24 (1.8)	2 (0.5)
Minnesota	83 (1.9)	16 (1.9)	1 (0.5)
Mississippi	51 (2.0)	47 (2.0)	3 (0.7)
Missouri	78 (2.0)	20 (1.9)	2 (0.5)
Nebraska	82 (1.4)	17 (1.4)	1 (0.4)
New Hampshire	84 (1.7)	15 (1.6)	1 (0.5)
New Jersey	72 (2.2)	27 (2.1)	1 (0.6)
New Mexico	69 (1.7)	26 (1.7)	4 (1.0)
New York	73 (2.7)	23 (2.3)	4 (1.5)
North Carolina	67 (2.1)	31 (2.2)	2 (0.5)
North Dakota	84 (1.8)	16 (1.8)	0 (0.2)
Ohio	71 (2.8)	27 (2.7)	2 (0.6)
Oklahoma	74 (1.9)	23 (1.9)	2 (0.7)
Pennsylvania	81 (1.7)	18 (1.7)	1 (0.4)
Rhode Island	76 (2.3)	23 (2.2)	2 (0.6)
South Carolina	70 (1.8)	28 (1.6)	2 (0.5)
Tennessee	69 (2.2)	29 (2.1)	1 (0.5)
Texas	70 (2.3)	28 (2.3)	2 (0.4)
Utah	81 (1.9)	18 (1.9)	1 (0.4)
Virginia	72 (1.8)	26 (1.8)	1 (0.3)
West Virginia	71 (2.2)	27 (2.3)	2 (0.6)
Wisconsin	81 (1.4)	19 (1.3)	1 (0.3)
Wyoming	82 (1.7)	17 (1.7)	1 (0.4)
TERRITORIES			
Guam	55 (2.5)	40 (2.8)	5 (1.3)
Virgin Islands	38 (2.4)	46 (3.0)	15 (2.1)

TABLE 1.15 | **Percentages of Correct Responses to Regular Constructed-Response Questions "Using Ruler to Measure Centimeters" (continued)**

PUBLIC SCHOOLS	Grade 4 - 1992		
	Measure Diagnoal		
	Correct	Incorrect	No Response
NATION	59 (1.3)	34 (1.2)	7 (0.8)
Northeast	66 (3.9)	27 (2.5)	7 (2.3)
Southeast	50 (2.8)	45 (2.1)	5 (1.4)
Central	67 (2.9)	28 (3.2)	5 (1.5)
West	56 (2.1)	34 (1.7)	10 (1.5)
STATES			
Alabama	51 (2.1)	45 (2.2)	4 (0.9)
Arizona	56 (2.0)	39 (2.0)	5 (0.8)
Arkansas	51 (1.8)	43 (1.8)	6 (1.1)
California	51 (2.3)	39 (2.1)	10 (1.3)
Colorado	62 (2.1)	34 (1.9)	4 (0.7)
Connecticut	67 (2.2)	30 (2.2)	4 (0.8)
Delaware	59 (2.1)	38 (2.1)	4 (1.0)
Dist. Columbia	28 (2.2)	60 (2.4)	13 (1.6)
Florida	54 (2.8)	40 (3.1)	6 (1.1)
Georgia	52 (2.3)	44 (2.2)	4 (0.9)
Hawaii	56 (2.4)	39 (2.1)	6 (1.2)
Idaho	62 (1.9)	33 (1.7)	5 (0.9)
Indiana	61 (2.1)	37 (2.0)	2 (0.7)
Iowa	68 (1.5)	30 (1.5)	2 (0.6)
Kentucky	51 (1.9)	45 (1.8)	4 (0.7)
Louisiana	47 (2.5)	48 (2.3)	5 (0.8)
Maine	65 (2.3)	32 (2.5)	3 (1.0)
Maryland	57 (1.9)	39 (1.9)	4 (1.1)
Massachusetts	64 (2.4)	32 (2.1)	5 (1.2)
Michigan	62 (2.3)	34 (2.2)	4 (0.9)
Minnesota	66 (2.0)	30 (1.9)	4 (0.7)
Mississippi	43 (2.2)	52 (2.1)	5 (0.8)
Missouri	65 (1.9)	32 (2.0)	3 (0.8)
Nebraska	66 (2.7)	32 (2.7)	2 (0.5)
New Hampshire	68 (2.1)	27 (1.8)	5 (0.8)
New Jersey	62 (1.9)	34 (2.0)	4 (1.1)
New Mexico	52 (2.9)	44 (2.4)	5 (1.1)
New York	54 (2.8)	40 (2.7)	7 (1.3)
North Carolina	51 (2.2)	43 (2.0)	6 (1.0)
North Dakota	70 (2.2)	28 (2.0)	2 (0.7)
Ohio	59 (2.4)	37 (2.2)	4 (0.8)
Oklahoma	60 (2.7)	37 (2.5)	3 (0.8)
Pennsylvania	64 (1.9)	33 (1.8)	3 (0.7)
Rhode Island	55 (2.0)	37 (2.1)	8 (1.5)
South Carolina	53 (2.2)	44 (2.1)	4 (0.8)
Tennessee	49 (2.2)	47 (1.9)	4 (1.0)
Texas	56 (2.4)	40 (2.3)	3 (0.7)
Utah	66 (2.1)	31 (2.1)	4 (0.8)
Virginia	59 (2.5)	38 (2.4)	3 (1.0)
West Virginia	57 (2.1)	40 (2.1)	3 (0.7)
Wisconsin	68 (2.4)	31 (2.4)	2 (0.5)
Wyoming	65 (2.2)	30 (2.0)	5 (0.9)
TERRITORY			
Guam	34 (2.0)	56 (2.5)	10 (1.8)

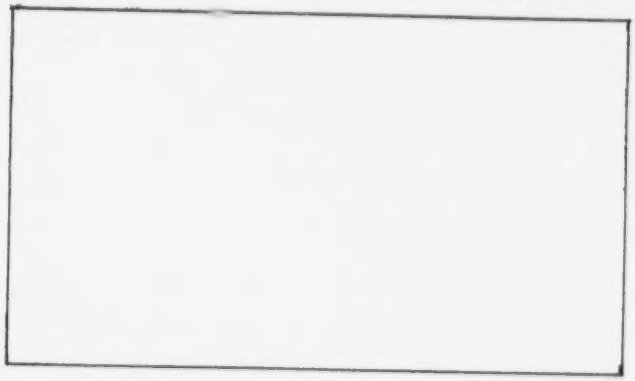
TABLE 1.15 | **Percentages of Correct Responses to Regular Constructed-Response Questions "Using Ruler to Measure Centimeters" (continued)**

PUBLIC SCHOOLS	Grade 8 - 1992		
	Measure Diagnoal		
	Correct	Incorrect	No Response
NATION	77 (1.2)	19 (1.2)	3 (0.6)
Northeast	81 (2.7)	16 (2.7)	3 (0.8)
Southeast	74 (2.5)	23 (1.5)	3 (1.3)
Central	80 (2.2)	17 (2.2)	3 (1.5)
West	75 (2.5)	21 (2.8)	4 (0.8)
STATES			
Alabama	75 (2.2)	23 (1.9)	2 (0.8)
Arizona	82 (1.9)	15 (1.6)	3 (0.7)
Arkansas	78 (2.0)	21 (1.9)	1 (0.5)
California	76 (1.9)	20 (1.6)	4 (0.8)
Colorado	84 (1.3)	14 (1.2)	2 (0.5)
Connecticut	87 (1.7)	12 (1.7)	1 (0.6)
Delaware	80 (1.9)	17 (1.7)	3 (0.8)
Dist. Columbia	60 (2.1)	31 (2.2)	9 (1.4)
Florida	76 (2.2)	21 (2.1)	3 (0.8)
Georgia	78 (2.2)	20 (2.1)	2 (0.7)
Hawaii	78 (1.7)	19 (1.6)	4 (0.7)
Idaho	86 (1.6)	12 (1.5)	2 (0.6)
Indiana	84 (1.6)	15 (1.6)	1 (0.4)
Iowa	88 (1.5)	11 (1.3)	1 (0.4)
Kentucky	81 (1.7)	17 (1.6)	2 (0.5)
Louisiana	74 (2.1)	22 (2.2)	4 (0.7)
Maine	87 (1.5)	12 (1.4)	1 (0.5)
Maryland	76 (1.7)	19 (1.8)	4 (1.0)
Massachusetts	85 (1.6)	12 (1.3)	3 (0.9)
Michigan	79 (1.7)	19 (1.6)	2 (0.5)
Minnesota	87 (1.2)	12 (1.2)	1 (0.3)
Mississippi	67 (2.1)	31 (2.0)	2 (0.5)
Missouri	85 (1.7)	13 (1.5)	2 (0.5)
Nebraska	85 (1.6)	14 (1.7)	1 (0.5)
New Hampshire	90 (1.3)	10 (1.2)	0 (0.3)
New Jersey	83 (2.1)	16 (2.0)	1 (0.5)
New Mexico	78 (1.5)	18 (1.3)	4 (0.9)
New York	82 (2.4)	13 (2.0)	5 (1.6)
North Carolina	80 (1.7)	19 (1.7)	2 (0.4)
North Dakota	88 (1.6)	11 (1.6)	0 (0.2)
Ohio	80 (2.3)	18 (2.2)	2 (0.8)
Oklahoma	82 (2.0)	16 (2.0)	2 (0.7)
Pennsylvania	89 (1.4)	11 (1.4)	1 (0.4)
Rhode Island	87 (1.5)	11 (1.4)	1 (0.5)
South Carolina	80 (1.8)	18 (1.7)	2 (0.5)
Tennessee	78 (1.9)	20 (1.8)	1 (0.5)
Texas	77 (1.9)	22 (1.8)	2 (0.4)
Utah	85 (1.4)	14 (1.5)	1 (0.4)
Virginia	83 (1.7)	16 (1.7)	1 (0.4)
West Virginia	81 (2.0)	17 (2.0)	2 (0.6)
Wisconsin	85 (1.5)	14 (1.5)	1 (0.3)
Wyoming	86 (1.4)	13 (1.3)	1 (0.5)
TERRITORIES			
Guam	69 (2.5)	26 (2.8)	5 (1.2)
Virgin Islands	53 (2.2)	32 (2.7)	14 (2.1)

The following geometry question involving the use of a ruler required students to draw a rectangle 2 inches wide by 3½ inches long. Students were provided space in which they could draw the rectangle in any orientation, horizontal, vertical, or otherwise. To receive credit for their drawing, the rectangle needed reasonably straight sides, vertex angles closely approximating 90°, and sides satisfying the dimensions of **2 ± ⅜** inches in width and **3½ ± ⅜** inches in length. To evaluate the students' responses, a template was constructed for the scorers to place above a student constructed rectangle to see if it fell within the boundaries defined by the above conditions. This provided a uniform grading procedure for responses to this task. Nationally, 18 percent of grade 4 students and 58 percent of grade 8 students correctly completed the task.

EXAMPLE 10: Geometry

In the space below, draw a rectangle 2 inches wide and $3\frac{1}{2}$ inches long.



Overall Percent Correct*
Grade 4 -- 18 (0.9)
Grade 8 -- 58 (1.3)

*The standard errors of the estimated percentages appear in parentheses.

As can be seen in TABLE 1.16, the substantial increase in percentage of correct responses between grades 4 and 8 noted for the nation also occurred for each of the various subgroups. At grade 4, from 8 to 27 percent of students, depending on the subgroup, did not provide a response to the question. This may have been due to the use of fractions in the measurements. At grade 8, most of the students attempted to provide a response to this task.

TABLE 1.16 National Results for Demographic Subgroups for the Regular Constructed-Response Task, "Using Ruler to Draw a Rectangle"

Grade 4

	Correct	Incorrect	No Response
Nation	18 (0.9)	68 (1.2)	13 (0.9)
Northeast	16 (2.3)	76 (3.0)	8 (2.0)
Southeast	17 (1.4)	71 (2.4)	12 (1.6)
Central	20 (2.2)	69 (2.7)	11 (1.7)
West	20 (1.7)	61 (1.6)	19 (1.5)
White	22 (1.3)	67 (1.5)	11 (1.0)
Black	6 (1.4)	73 (3.3)	21 (3.3)
Hispanic	11 (1.8)	70 (3.0)	19 (2.7)
Male	18 (1.4)	64 (1.7)	18 (1.3)
Female	18 (1.6)	74 (1.8)	8 (1.1)
Advantaged Urban	27 (3.0)	64 (3.0)	9 (2.4)
Disadvantaged Urban	8 (1.9)	65 (4.8)	27 (4.7)
Extreme Rural	21 (6.0)	68 (6.0)	11 (3.0)
Other	18 (1.2)	70 (1.6)	12 (1.0)
Public	17 (1.1)	69 (1.4)	14 (1.0)
Catholic and Other Private	29 (2.5)	62 (2.9)	9 (1.8)

Grade 8

	Correct	Incorrect	No Response
Nation	58 (1.3)	38 (1.2)	4 (0.5)
Northeast	63 (3.0)	34 (2.7)	3 (1.0)
Southeast	52 (2.9)	44 (2.6)	4 (1.5)
Central	65 (2.6)	34 (2.6)	1 (0.6)
West	54 (2.0)	40 (1.6)	5 (1.0)
White	68 (1.5)	30 (1.4)	2 (0.3)
Black	28 (3.4)	65 (4.1)	7 (2.3)
Hispanic	37 (2.9)	55 (3.1)	8 (2.0)
Male	59 (2.0)	36 (1.9)	5 (0.9)
Female	58 (1.8)	40 (1.8)	2 (0.6)
Advantaged Urban	71 (2.7)	29 (2.7)	0 (0.0)
Disadvantaged Urban	34 (2.9)	58 (3.2)	8 (2.1)
Extreme Rural	56 (5.9)	42 (5.6)	2 (1.1)
Other	60 (1.6)	37 (1.5)	4 (0.7)
Public	57 (1.4)	39 (1.3)	4 (0.6)
Catholic and Other Private	67 (2.5)	31 (2.6)	2 (0.7)

The standard errors of the estimated percentages appear in parentheses. It can be said with about 95 percent certainty that for each population of interest, the value for the whole populations is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details. When the proportion of students is either 0 percent or 100 percent, the standard error is inestimable. However, percentages 99.5 percent and greater were rounded to 100 percent and percentages 0.5 percent or less were rounded to 0 percent. Percentages may not total 100 percent due to rounding error.

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment

TABLE 1.17 contains the state-by-state data for responses to this task by fourth- and eighth-grade students. Again, considerable variation existed between states -- from 8 to 31 percent correct at grade 4 and from 30 to 77 percent correct at grade 8. Yet, the increase in performance between grades 4 and 8 held both within and across states.

TABLE 1.17 | **Percentage of Correct Responses to Regular Constructed-Response Question, "Using Ruler to Draw a Rectangle"**

PUBLIC SCHOOLS	Grade 4 - 1992		
	Correct	Incorrect	No Response
NATION	17 (1.1)	69 (1.4)	14 (1.0)
Northeast	14 (1.9)	76 (3.3)	9 (2.4)
Southeast	16 (1.3)	72 (2.4)	12 (1.7)
Central	17 (3.0)	71 (3.5)	12 (2.0)
West	19 (1.8)	61 (1.6)	20 (1.8)
STATES			
Alabama	12 (1.5)	72 (1.8)	15 (1.5)
Arizona	20 (1.8)	66 (2.1)	15 (1.6)
Arkansas	12 (1.5)	74 (2.4)	14 (1.8)
California	18 (1.7)	63 (2.3)	20 (1.8)
Colorado	25 (1.8)	65 (1.8)	10 (1.1)
Connecticut	26 (1.9)	63 (2.2)	11 (1.4)
Delaware	17 (1.5)	71 (1.7)	12 (1.6)
Dist. Columbia	8 (1.1)	70 (1.8)	22 (1.7)
Florida	16 (2.5)	68 (2.7)	16 (1.7)
Georgia	16 (1.7)	71 (1.9)	13 (1.5)
Hawaii	20 (1.6)	68 (1.9)	12 (1.2)
Idaho	24 (1.7)	65 (2.2)	11 (1.5)
Indiana	21 (1.9)	69 (2.0)	10 (1.3)
Iowa	28 (1.8)	64 (1.9)	8 (1.1)
Kentucky	16 (1.5)	73 (1.7)	11 (1.2)
Louisiana	16 (1.6)	72 (1.7)	12 (1.3)
Maine	31 (2.5)	66 (2.5)	4 (0.8)
Maryland	21 (1.4)	67 (1.5)	12 (1.3)
Massachusetts	22 (1.8)	69 (2.1)	9 (1.1)
Michigan	19 (1.6)	68 (1.8)	13 (1.4)
Minnesota	26 (1.8)	66 (2.1)	8 (1.0)
Mississippi	12 (1.4)	74 (1.9)	14 (1.7)
Missouri	21 (1.6)	71 (1.7)	8 (1.1)
Nebraska	28 (2.6)	65 (2.4)	7 (1.0)
New Hampshire	27 (2.1)	64 (1.9)	9 (1.3)
New Jersey	25 (2.2)	67 (2.1)	8 (1.2)
New Mexico	21 (2.4)	69 (2.7)	10 (1.3)
New York	19 (1.7)	68 (2.1)	13 (1.8)
North Carolina	16 (1.4)	70 (1.8)	14 (1.4)
North Dakota	30 (2.5)	63 (2.2)	6 (1.0)
Ohio	22 (1.8)	67 (2.1)	12 (1.7)
Oklahoma	15 (1.7)	78 (1.9)	7 (1.1)
Pennsylvania	22 (1.7)	67 (2.1)	11 (1.3)
Rhode Island	18 (2.1)	70 (2.4)	12 (1.6)
South Carolina	14 (1.5)	76 (1.6)	10 (1.3)
Tennessee	13 (1.6)	71 (2.0)	16 (1.8)
Texas	20 (1.8)	72 (2.2)	7 (1.1)
Utah	28 (2.4)	61 (2.3)	11 (1.3)
Virginia	22 (2.1)	69 (2.1)	9 (1.1)
West Virginia	17 (1.6)	72 (1.9)	11 (1.3)
[illegible]sin	26 (2.1)	68 (2.3)	6 (0.9)
[illegible]g	24 (2.0)	69 (2.0)	7 (0.9)
[illegible]ORY			
[illegible]	8 (1.1)	76 (2.2)	16 (1.8)

The standard errors of the estimated percentages appear in parentheses. It can be said with 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details). Percentages may not total 100 percent due to rounding error.

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment.

TABLE 1.17 | **Percentage of Correct Responses to Regular Constructed-Response Question, "Using Ruler to Draw a Rectangle" (continued)**

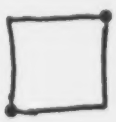
PUBLIC SCHOOLS	Grade 8 - 1992		
	Correct	Incorrect	No Response
NATION	57 (1.4)	39 (1.3)	4 (0.6)
Northeast	62 (3.6)	34 (3.2)	4 (1.2)
Southeast	49 (3.1)	46 (2.8)	4 (1.6)
Central	64 (2.6)	35 (2.7)	1 (0.6)
West	54 (2.3)	41 (1.9)	5 (1.1)
STATES			
Alabama	50 (2.3)	48 (2.2)	3 (0.8)
Arizona	62 (1.7)	34 (1.9)	4 (1.0)
Arkansas	51 (2.3)	46 (2.3)	3 (0.8)
California	61 (2.4)	35 (2.0)	4 (1.0)
Colorado	66 (2.2)	31 (2.2)	3 (0.6)
Connecticut	69 (2.0)	28 (2.3)	3 (0.8)
Delaware	60 (2.8)	35 (2.7)	5 (1.3)
Dist. Columbia	35 (2.7)	53 (2.5)	12 (1.8)
Florida	56 (2.4)	40 (2.5)	4 (0.9)
Georgia	49 (2.3)	47 (2.4)	3 (0.8)
Hawaii	60 (2.4)	36 (2.6)	3 (0.8)
Idaho	69 (1.9)	29 (1.9)	2 (0.5)
Indiana	66 (2.3)	32 (2.2)	2 (0.5)
Iowa	71 (2.0)	27 (2.0)	2 (0.6)
Kentucky	60 (2.2)	36 (2.0)	4 (1.0)
Louisiana	40 (2.3)	55 (2.3)	5 (0.7)
Maine	73 (1.6)	27 (1.6)	1 (0.4)
Maryland	57 (2.0)	39 (2.0)	4 (1.0)
Massachusetts	66 (1.8)	31 (1.9)	3 (0.7)
Michigan	60 (2.3)	36 (2.3)	4 (0.8)
Minnesota	73 (2.1)	26 (1.9)	1 (0.4)
Mississippi	40 (1.8)	55 (2.1)	5 (1.1)
Missouri	66 (2.1)	32 (2.2)	2 (0.5)
Nebraska	73 (1.7)	25 (1.7)	2 (0.6)
New Hampshire	71 (1.9)	27 (1.7)	2 (0.7)
New Jersey	58 (2.5)	39 (2.6)	3 (0.8)
New Mexico	58 (2.0)	39 (2.0)	3 (0.8)
New York	59 (2.5)	35 (2.4)	6 (1.3)
North Carolina	55 (1.8)	41 (1.8)	4 (0.8)
North Dakota	77 (1.8)	22 (1.8)	0 (0.2)
Ohio	62 (2.4)	35 (2.3)	3 (0.9)
Oklahoma	65 (2.3)	33 (2.3)	3 (0.7)
Pennsylvania	69 (1.9)	29 (1.8)	3 (0.7)
Rhode Island	57 (2.6)	41 (2.7)	3 (0.7)
South Carolina	53 (2.4)	45 (2.3)	2 (0.5)
Tennessee	53 (2.5)	44 (2.5)	3 (0.8)
Texas	57 (2.5)	40 (2.5)	3 (0.8)
Utah	71 (1.9)	28 (1.8)	1 (0.5)
Virginia	61 (2.3)	38 (2.2)	2 (0.5)
West Virginia	56 (1.8)	40 (1.8)	4 (0.8)
Wisconsin	72 (2.2)	26 (2.2)	2 (0.6)
Wyoming	73 (2.3)	25 (2.0)	2 (0.6)
TERRITORIES			
Guam	46 (3.0)	45 (2.9)	9 (1.7)
Virgin Islands	30 (2.8)	59 (3.0)	11 (1.4)

Another task provided students with two dots and asked them to use their rulers to draw a square with two of its corners at the points shown. As can be seen below, this geometry question required students to use their spatial perception skills. They needed to be able to visualize at least one of the three potential squares that could satisfy the conditions defined by the two corner points. The first and second answers were squares developed under the assumption that the points were adjacent vertices of the square. This leads to squares drawn either above or below the segment joining the two points. The third solution resulted from seeing the two points as opposite ends of a diagonal of the desired square. In this case, the sides of the square were basically parallel to the sides of the pages of the test booklet. This question, like the earlier regular constructed-response item involving the sections of the cylinder, assessed aspects of students' spatial perception.

EXAMPLE 11: Geometry

In the space below, use your ruler to draw a square with two of its corners at the points shown.

Overall Percent Correct *
Grade 4 -- 40 (1.3)
Grade 8 -- 67 (1.5)



(One of the possible answers)

*The standard errors of the estimated percentages appear in parentheses.

Student responses to this task were scored using a template, similar to the previous question. The major criterion used in developing the template was that at least three sides were equal in length and the figure closely approximated a square in shape. Nationally, 40 percent of the fourth graders and 67 percent of the eighth graders correctly constructed a solution to this task. TABLES 1.18 and 1.19 contain the information showing subgroup and state-level student responses to this task. At grade 4, state performance ranged from 37 to 52 percent correct. At grade 8, the corresponding range of performance was from 45 to 81 percent correct.

TABLE 1.18 National Results for Demographic Subgroups for the Regular Constructed-Response Task, "Using Ruler to Draw a Square"

Grade 4

	Correct	Incorrect	No Response
Nation	40 (1.3)	52 (1.2)	8 (0.6)
Northeast	44 (2.9)	50 (2.9)	5 (1.5)
Southeast	30 (1.5)	61 (1.9)	9 (1.5)
Central	43 (3.2)	50 (2.9)	8 (0.9)
West	45 (2.0)	48 (2.2)	7 (1.2)
White	47 (1.7)	48 (1.7)	5 (0.6)
Black	15 (2.7)	68 (3.6)	18 (3.0)
Hispanic	37 (3.1)	55 (3.3)	8 (1.6)
Male	41 (2.3)	50 (2.2)	9 (1.1)
Female	40 (1.9)	54 (2.1)	6 (0.8)
Advantaged Urban	51 (5.2)	46 (4.8)	3 (1.4)
Disadvantaged Urban	18 (3.1)	58 (4.2)	24 (4.1)
Extreme Rural	46 (6.7)	50 (6.7)	4 (1.2)
Other	41 (1.5)	53 (1.7)	7 (0.7)
Public	39 (1.4)	53 (1.4)	8 (0.7)
Catholic and Other Private	48 (3.1)	48 (2.9)	4 (1.1)

Grade 8

	Correct	Incorrect	No Response
Nation	67 (1.5)	32 (1.7)	2 (0.5)
Northeast	69 (2.9)	30 (2.9)	0 (0.3)
Southeast	59 (2.2)	38 (3.1)	3 (1.6)
Central	71 (1.9)	27 (1.9)	1 (0.6)
West	68 (4.4)	30 (4.6)	2 (0.6)
White	73 (1.9)	26 (2.0)	1 (0.3)
Black	43 (2.9)	52 (3.6)	5 (2.4)
Hispanic	62 (3.8)	37 (3.8)	2 (0.6)
Male	69 (1.7)	28 (1.9)	3 (0.8)
Female	64 (2.1)	35 (2.3)	1 (0.5)
Advantaged Urban	77 (2.2)	22 (2.2)	0 (0.2)
Disadvantaged Urban	54 (3.0)	38 (4.4)	8 (3.6)
Extreme Rural	70 (5.7)	28 (5.8)	2 (1.1)
Other	66 (2.0)	32 (2.1)	1 (0.4)
Public	66 (1.7)	32 (1.9)	2 (0.5)
Catholic and Other Private	75 (2.2)	24 (2.1)	1 (0.6)

The standard errors of the estimated percentages appear in parentheses. It can be said with about 95 percent certainty that for each population of interest, the value for the whole populations is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details). When the proportion of students is either 0 percent or 100 percent, the standard error is inestimable. However, percentages 99.5 percent and greater were rounded to 100 percent and percentages 0.5 percent or less were rounded to 0 percent. Percentages may not total 100 percent due to rounding error.

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment

TABLE 1.19 | **Percentage of Correct Responses to Regular Constructed-Response Question, "Using Ruler to Draw a Square"**

PUBLIC SCHOOLS	Grade 4 - 1992		
	Correct	Incorrect	No Response
NATION	39 (1.4)	53 (1.4)	8 (0.7)
Northeast	46 (3.8)	49 (3.8)	6 (1.7)
Southeast	28 (1.5)	63 (2.0)	9 (1.7)
Central	41 (3.4)	51 (3.3)	9 (1.0)
West	45 (2.2)	48 (2.5)	7 (1.3)
STATES			
Alabama	33 (2.5)	60 (2.3)	7 (1.2)
Arizona	43 (1.9)	49 (2.0)	8 (1.1)
Arkansas	35 (2.3)	56 (2.5)	8 (1.2)
California	33 (2.2)	52 (2.5)	15 (1.5)
Colorado	43 (1.9)	51 (2.1)	6 (1.0)
Connecticut	45 (2.4)	49 (2.3)	6 (0.9)
Delaware	36 (2.2)	55 (2.5)	9 (1.4)
Dist. Columbia	22 (2.1)	62 (2.2)	16 (1.4)
Florida	37 (2.5)	55 (2.2)	9 (1.4)
Georgia	35 (2.0)	57 (2.0)	8 (1.2)
Hawaii	38 (2.5)	53 (2.4)	8 (1.3)
Idaho	46 (2.0)	46 (2.1)	8 (1.3)
Indiana	39 (2.4)	54 (2.4)	7 (1.2)
Iowa	49 (2.3)	45 (2.2)	5 (1.0)
Kentucky	35 (2.0)	58 (2.0)	7 (1.0)
Louisiana	33 (2.0)	58 (2.1)	10 (1.3)
Maine	52 (2.5)	43 (2.3)	5 (1.0)
Maryland	39 (2.0)	54 (2.4)	8 (1.0)
Massachusetts	48 (2.3)	47 (2.4)	6 (0.9)
Michigan	41 (2.5)	50 (2.3)	9 (1.3)
Minnesota	50 (2.3)	45 (2.2)	5 (0.9)
Mississippi	27 (2.1)	62 (2.1)	11 (1.5)
Missouri	48 (2.1)	47 (2.1)	5 (0.9)
Nebraska	44 (2.8)	51 (2.6)	5 (1.1)
New Hampshire	50 (2.7)	44 (2.5)	6 (1.1)
New Jersey	44 (2.4)	50 (2.4)	6 (1.2)
New Mexico	41 (2.0)	53 (2.3)	6 (1.4)
New York	41 (2.3)	51 (2.4)	8 (1.1)
North Carolina	33 (2.2)	61 (2.3)	7 (1.1)
North Dakota	52 (2.6)	45 (2.6)	4 (1.1)
Ohio	43 (2.3)	50 (2.2)	7 (0.9)
Oklahoma	45 (1.9)	51 (2.0)	5 (0.9)
Pennsylvania	42 (2.7)	52 (2.8)	5 (1.0)
Rhode Island	40 (2.9)	52 (2.9)	9 (1.8)
South Carolina	32 (1.7)	62 (1.8)	7 (1.2)
Tennessee	30 (1.8)	62 (2.1)	9 (1.5)
Texas	40 (1.9)	56 (2.0)	3 (0.8)
Utah	49 (2.2)	45 (2.0)	6 (1.0)
Virginia	39 (2.2)	56 (2.1)	5 (1.0)
West Virginia	39 (2.1)	55 (2.2)	6 (0.9)
Wisconsin	46 (1.9)	49 (2.0)	5 (1.1)
Wyoming	45 (2.1)	50 (2.1)	5 (1.0)
TERRITORY			
Guam	37 (2.2)	52 (2.0)	11 (1.7)

The standard errors of the estimated percentages appear in parentheses. It can be said with 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details). Percentages may not total 100 percent due to rounding error.

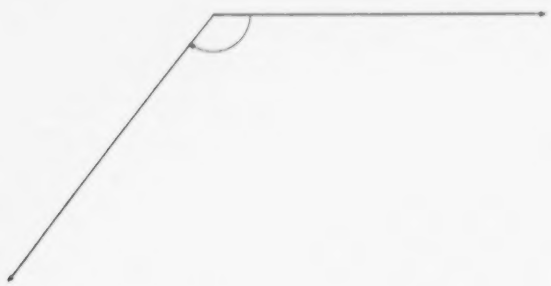
SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment.

TABLE 1.19 | **Percentage of Correct Responses to Regular Constructed-Response Question, "Using Ruler to Draw a Square" (continued)**

PUBLIC SCHOOLS	Grade 8 - 1992		
	Correct	Incorrect	No Response
NATION	66 (1.7)	32 (1.9)	2 (0.5)
Northeast	67 (3.5)	33 (3.4)	0 (0.3)
Southeast	57 (2.3)	39 (3.6)	3 (1.8)
Central	72 (2.1)	27 (2.1)	1 (0.7)
West	67 (4.8)	31 (5.0)	2 (0.6)
STATES			
Alabama	61 (2.2)	37 (2.1)	2 (0.6)
Arizona	74 (1.8)	24 (1.7)	2 (0.5)
Arkansas	65 (2.1)	33 (2.1)	2 (0.6)
California	70 (2.3)	28 (2.2)	2 (0.6)
Colorado	79 (1.6)	19 (1.5)	1 (0.4)
Connecticut	74 (2.3)	25 (2.3)	1 (0.4)
Delaware	68 (2.5)	31 (2.5)	1 (0.6)
Dist. Columbia	48 (2.6)	47 (2.5)	5 (1.0)
Florida	63 (2.3)	36 (2.3)	1 (0.4)
Georgia	63 (2.1)	36 (2.1)	2 (0.4)
Hawaii	71 (2.3)	27 (2.2)	2 (0.6)
Idaho	78 (1.8)	21 (1.9)	1 (0.3)
Indiana	73 (2.2)	26 (2.2)	1 (0.4)
Iowa	78 (1.8)	22 (1.8)	0 (0.3)
Kentucky	70 (2.0)	29 (1.9)	2 (0.6)
Louisiana	56 (2.3)	43 (2.3)	1 (0.3)
Maine	78 (2.2)	21 (2.2)	1 (0.4)
Maryland	70 (2.3)	29 (2.3)	1 (0.4)
Massachusetts	73 (2.1)	26 (2.1)	1 (0.3)
Michigan	69 (2.3)	30 (2.3)	2 (0.5)
Minnesota	77 (2.0)	22 (2.1)	1 (0.4)
Mississippi	53 (2.1)	46 (2.2)	1 (0.4)
Missouri	71 (1.8)	28 (1.7)	1 (0.4)
Nebraska	78 (2.0)	21 (2.0)	1 (0.5)
New Hampshire	81 (1.9)	19 (1.8)	0 (0.3)
New Jersey	71 (2.1)	28 (1.9)	1 (0.5)
New Mexico	73 (2.2)	25 (2.2)	1 (0.5)
New York	64 (2.6)	35 (2.7)	2 (0.8)
North Carolina	64 (2.1)	35 (2.1)	2 (0.4)
North Dakota	78 (2.3)	22 (2.3)	0 (0.0)
Ohio	73 (2.3)	26 (2.2)	1 (0.6)
Oklahoma	73 (1.9)	26 (2.0)	1 (0.5)
Pennsylvania	73 (2.1)	25 (1.9)	2 (0.7)
Rhode Island	69 (2.4)	29 (2.4)	2 (0.4)
South Carolina	62 (2.4)	37 (2.5)	2 (0.6)
Tennessee	63 (2.4)	35 (2.3)	1 (0.5)
Texas	66 (2.2)	32 (2.1)	2 (0.6)
Utah	80 (1.7)	19 (1.6)	1 (0.3)
Virginia	69 (2.1)	30 (2.1)	1 (0.5)
West Virginia	71 (1.8)	28 (1.9)	1 (0.5)
Wisconsin	73 (1.7)	26 (1.7)	1 (0.2)
Wyoming	73 (2.4)	26 (2.4)	1 (0.5)
TERRITORIES			
Guam	63 (3.0)	32 (3.0)	6 (1.4)
Virgin Islands	45 (2.8)	46 (2.7)	10 (2.0)

The final example of a constructed-response task involving a ruler/protractor was a measurement question asking grade 8 students to use their protractors to find the degree measure of an angle showing a negative rotation of approximately 127° from standard position. Given the orientation of the angle and consideration of sample student performance in the field testing and the early responses on the actual NAEP, it was decided to award credit for any answer in the range 124° to 130°. Since the question asked students to find the degree measure, students were not required to write the degree symbol with their numerical answer. Also, the lack of precision created by the cardboard protractor provided to students, which was graduated only to whole degrees, in combination with some amount of imprecision as a result of the diagram, meant that students were never expected to provide answers to more than the nearest whole degree of accuracy.

EXAMPLE 12: Measurement



Overall Percent Correct *
Grade 8 -- 35 (1.9)

Use your protractor to find the degree measure of the angle shown above.

Answer: 127°

*The standard errors of the estimated percentages appear in parentheses.

Nationally, 35 percent of grade 8 students completed this task correctly, and percentages of success were significantly lower for several subgroups (see TABLE 1.20). TABLE 1.21 contains the results for state-by-state student performance. The percentages of correct performance ranged from 12 to 53 percent, further indicating the low level of skills that students leaving middle school have in the reading and application of rulers and protractors.

TABLE 1.20 National Results for Demographic Subgroups for the Regular Constructed-Response Task, "Using Protractor to Measure an Angle"

Grade 8

	Correct	Incorrect	No Response
Nation	35 (1.9)	55 (1.7)	10 (1.0)
Northeast	41 (6.2)	50 (4.8)	9 (2.4)
Southeast	28 (3.2)	61 (3.2)	11 (2.1)
Central	40 (3.7)	54 (3.5)	6 (1.8)
West	32 (2.9)	56 (1.9)	12 (2.0)
White	41 (2.4)	52 (2.0)	7 (0.9)
Black	15 (3.7)	68 (4.9)	17 (3.3)
Hispanic	18 (2.9)	62 (3.9)	20 (3.3)
Male	37 (2.8)	55 (2.5)	8 (1.2)
Female	33 (1.8)	56 (1.7)	11 (1.4)
Advantaged Urban	42 (4.1)	53 (3.4)	5 (1.7)
Disadvantaged Urban	17 (5.1)	66 (5.5)	17 (4.0)
Extreme Rural	34 (5.9)	57 (5.0)	9 (3.5)
Other	36 (2.3)	54 (1.9)	10 (1.4)
Public	35 (2.1)	56 (1.8)	10 (1.1)
Catholic and Other Private	38 (3.1)	53 (2.7)	9 (1.3)

The standard errors of the estimated percentages appear in parentheses. It can be said with about 95 percent certainty that for each population of interest, the value for the whole populations within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details). Percentages may not total 100 percent due to rounding error.

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment

TABLE 1.21 | **Percentage of Correct Responses to Regular Constructed-Response Question, "Using Protractor to Measure an Angle"**

PUBLIC SCHOOLS	Grade 8 - 1992		
	Correct	Incorrect	No Response
NATION	35 (2.1)	56 (1.8)	10 (1.1)
Northeast	42 (6.9)	50 (5.5)	9 (2.4)
Southeast	28 (3.5)	61 (3.6)	11 (2.3)
Central	39 (4.0)	55 (3.7)	6 (2.0)
West	32 (3.1)	56 (2.1)	12 (2.2)
STATES			
Alabama	24 (2.4)	65 (2.1)	10 (1.5)
Arizona	28 (2.5)	60 (2.4)	12 (1.7)
Arkansas	24 (2.2)	68 (2.3)	8 (1.1)
California	29 (2.5)	55 (2.2)	16 (1.7)
Colorado	37 (2.1)	55 (2.0)	8 (1.2)
Connecticut	41 (2.3)	52 (2.4)	7 (1.5)
Delaware	34 (2.6)	57 (2.6)	9 (1.8)
Dist. Columbia	20 (2.7)	57 (3.0)	23 (2.7)
Florida	29 (2.3)	59 (2.5)	12 (1.7)
Georgia	25 (2.4)	65 (2.3)	10 (1.5)
Hawaii	31 (2.6)	53 (2.6)	16 (2.0)
Idaho	43 (2.5)	48 (2.3)	9 (1.1)
Indiana	42 (2.5)	51 (2.4)	6 (1.0)
Iowa	53 (2.2)	45 (2.1)	2 (0.6)
Kentucky	32 (2.0)	59 (1.8)	8 (1.5)
Louisiana	22 (2.2)	69 (2.4)	10 (1.4)
Maine	42 (2.5)	51 (2.6)	6 (1.1)
Maryland	32 (2.1)	56 (2.6)	12 (1.8)
Massachusetts	37 (2.5)	53 (2.6)	10 (1.2)
Michigan	36 (2.7)	56 (2.5)	8 (1.2)
Minnesota	51 (2.9)	45 (2.7)	3 (0.8)
Mississippi	23 (2.7)	63 (2.5)	13 (1.9)
Missouri	45 (2.8)	51 (2.6)	4 (0.8)
Nebraska	46 (2.7)	51 (2.6)	3 (0.7)
New Hampshire	40 (2.8)	55 (2.6)	5 (1.2)
New Jersey	44 (2.3)	50 (2.2)	6 (1.2)
New Mexico	32 (2.2)	55 (2.2)	13 (1.2)
New York	42 (3.1)	49 (2.9)	10 (1.7)
North Carolina	30 (2.3)	61 (2.6)	8 (1.3)
North Dakota	52 (2.4)	45 (2.4)	3 (0.7)
Ohio	35 (3.7)	57 (3.5)	8 (1.3)
Oklahoma	26 (2.7)	68 (2.5)	6 (1.2)
Pennsylvania	41 (2.3)	52 (2.3)	6 (1.0)
Rhode Island	29 (2.8)	62 (3.0)	9 (1.7)
South Carolina	38 (2.3)	55 (2.5)	7 (1.2)
Tennessee	30 (2.6)	64 (2.7)	7 (1.3)
Texas	34 (2.8)	54 (2.3)	13 (1.6)
Utah	32 (2.2)	61 (2.4)	8 (1.2)
Virginia	36 (2.4)	56 (2.4)	8 (1.5)
West Virginia	28 (2.6)	62 (2.6)	10 (1.6)
Wisconsin	48 (2.5)	48 (2.2)	4 (0.9)
Wyoming	39 (2.3)	54 (2.2)	7 (1.0)
TERRITORIES			
Guam	26 (2.6)	54 (3.2)	20 (2.8)
Virgin Islands	12 (2.3)	53 (3.4)	35 (3.3)

The standard errors of the estimated percentages appear in parentheses. It can be said with 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details). Percentages may not total 100 percent due to rounding error.

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment.

Example Constructed-Response Questions with Calculator Available

The third group of constructed-response questions were those for which students had a hand calculator supplied. These questions were quite similar to the first set of constructed-response tasks presented, in that they usually required the construction of a numerical answer to be provided in a short-answer format. However, the tasks set by these questions usually called for a greater amount of calculation with more realistic numbers and data than do those a student sometimes finds in textbook-type application problems. For some of the calculator-assisted questions, students needed to investigate a situation and discover a pattern or a generalization which was based on a significant amount of numerical work.

Students in grade 4 were supplied with a Texas Instruments TI-108 calculator and students in grades 8 and 12 received a Texas Instruments TI-30 Challenger for their use. In addition, the students were provided with a short, prepared introduction to the use of these particular calculators prior to beginning the assessment exercises.

The two example calculator questions shown below were tasks presented only to fourth graders. The first of these asked students to examine four food items, which had calorie information given, and tell which two items would provide a total of approximately 600 calories. Correct answers could have been either cheeseburger and yogurt or some statement equivalent to two hotdogs. Forty-five percent of the fourth graders correctly completed this task.

The second question asked fourth graders to determine the amount of change that George would receive from a $10 bill if he purchased two calculators that cost $3.29 each. A student could receive credit for a correct response by responding $3.42, 3.42, 342 cents, or 342. Even under this liberal scoring for this numbers and operations question, only 21 percent of the fourth graders correctly answered the problem. The multi-step nature of the task may have contributed somewhat to the low level of performance.

EXAMPLE 13: Numbers and Operations

Cheeseburger
393 Calories

Hot Dog
298 Calories

Yogurt
214 Calories

Cookie
119 Calories

Overall Percent Correct*
Grade 4 – 45 (1.4)

Which two of the items above would provide a total of about 600 calories?

Answer: Cheeseburger
Yogurt

(One of the possible answers)

Did you use the calculator on this question?

(Yes) No

EXAMPLE 14: Numbers and Operations

George buys two calculators that cost $3.29 each. If there is no tax, how much change will he receive from a $10 bill?

Overall Percent Correct*
Grade 4 – 21 (1.4)

Answer: $3.42

Did you use the calculator on this question?

(Yes) No

*The standard errors of the estimated percentages appear in parentheses.

The results by subgroup further illustrate the difficulty fourth graders had with these questions, particularly the second one (see TABLE 1.22). The state data for these two numbers and operations questions, shown in TABLE 1.23, range from 28 to 55 percent correct for "Estimate Calories" and from 10 to 30 percent for "Change from Buying Two Calculators."

TABLE 1.22 National Results for Demographic Subgroups for the Regular Constructed-Response Tasks, "Estimate Calories" and "Change from Buying Two Calculators"

Grade 4

	Estimate Calories			Change from Buying Two Calculators		
	Correct	Incorrect	No Response	Correct	Incorrect	No Response
Nation	45 (1.4)	50 (1.5)	5 (0.7)	21 (1.4)	77 (1.6)	3 (0.5)
Northeast	47 (3.7)	47 (3.6)	6 (2.4)	26 (3.3)	73 (3.5)	1 (0.6)
Southeast	39 (3.3)	56 (3.7)	5 (1.1)	17 (2.4)	79 (2.5)	4 (1.6)
Central	50 (2.1)	46 (2.3)	4 (1.0)	20 (2.9)	78 (3.6)	2 (0.8)
West	45 (2.1)	49 (2.7)	6 (1.4)	20 (2.9)	76 (3.0)	4 (1.2)
White	52 (1.7)	44 (2.0)	5 (0.9)	23 (1.8)	75 (2.0)	2 (0.4)
Black	26 (2.6)	68 (2.8)	6 (1.5)	9 (2.3)	82 (3.6)	9 (2.8)
Hispanic	33 (3.5)	60 (3.6)	8 (1.7)	13 (2.3)	86 (2.3)	1 (0.5)
Male	44 (2.0)	52 (2.0)	4 (0.8)	20 (1.8)	78 (2.0)	3 (0.7)
Female	47 (2.0)	47 (2.1)	6 (0.9)	22 (2.2)	76 (2.4)	2 (0.9)
Advantaged Urban	54 (3.3)	44 (3.5)	3 (1.1)	26 (3.3)	74 (3.4)	1 (0.7)
Disadvantaged Urban	32 (4.6)	62 (4.8)	6 (2.5)	12 (3.4)	82 (4.6)	6 (2.7)
Extreme Rural	42 (3.7)	50 (4.8)	8 (3.3)	22 (3.7)	76 (3.2)	2 (1.4)
Other	46 (1.8)	49 (1.8)	5 (0.9)	20 (1.6)	77 (1.8)	3 (0.7)
Public	45 (1.6)	50 (1.7)	5 (0.8)	20 (1.6)	77 (1.8)	3 (0.5)
Catholic and Other Private	49 (3.6)	45 (3.7)	6 (1.4)	23 (3.3)	75 (3.1)	3 (2.1)

The standard errors of the estimated percentages appear in parentheses. It can be said with about 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details). Percentages may not total 100 percent due to rounding error.

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment

TABLE 1.23 | **Percentages of Correct Responses to Regular Constructed-Response Questions with Calculator Available, "Estimate Calories" and "Change from Buying Two Calculators"**

PUBLIC SCHOOLS	Grade 4 - 1992					
	Estimate Calories			Change from Buying Two Calculators		
	Correct	Incorrect	No Response	Correct	Incorrect	No Response
NATION	45 (1.6)	50 (1.7)	5 (0.8)	20 (1.6)	77 (1.8)	3 (0.5)
Northeast	45 (4.3)	48 (4.2)	7 (2.7)	26 (3.9)	73 (4.0)	1 (0.6)
Southeast	40 (3.7)	56 (4.0)	5 (1.0)	18 (2.6)	79 (2.6)	3 (1.1)
Central	50 (2.9)	47 (2.9)	3 (0.9)	19 (3.6)	80 (4.5)	2 (1.0)
West	45 (2.2)	50 (3.0)	5 (1.6)	20 (3.2)	76 (3.2)	4 (1.2)
STATES						
Alabama	38 (2.8)	59 (2.9)	3 (0.7)	16 (1.9)	84 (1.9)	1 (0.6)
Arizona	41 (2.1)	55 (2.2)	5 (0.9)	18 (1.9)	80 (2.0)	2 (0.7)
Arkansas	36 (2.0)	59 (2.1)	4 (0.9)	15 (1.7)	82 (1.7)	3 (0.8)
California	37 (2.5)	57 (2.5)	6 (1.0)	17 (2.0)	79 (1.9)	4 (0.9)
Colorado	47 (2.2)	49 (2.2)	4 (1.0)	20 (1.9)	79 (1.9)	1 (0.4)
Connecticut	49 (2.0)	45 (2.1)	5 (1.0)	23 (2.3)	76 (2.4)	1 (0.5)
Delaware	45 (2.9)	49 (2.7)	6 (1.1)	22 (2.2)	78 (2.1)	0 (0.3)
Dist. Columbia	28 (2.0)	64 (2.4)	8 (1.2)	12 (1.5)	84 (1.8)	4 (1.1)
Florida	42 (2.5)	51 (2.4)	7 (1.0)	17 (1.7)	81 (1.8)	2 (0.7)
Georgia	44 (2.2)	53 (2.2)	3 (0.8)	16 (1.8)	83 (2.1)	1 (0.8)
Hawaii	46 (2.3)	49 (2.4)	6 (1.0)	21 (1.8)	76 (2.0)	2 (0.8)
Idaho	46 (1.8)	49 (1.8)	5 (0.7)	19 (1.8)	80 (1.9)	0 (0.3)
Indiana	49 (2.4)	49 (2.3)	2 (0.6)	16 (1.4)	82 (1.5)	1 (0.6)
Iowa	51 (2.4)	45 (2.4)	4 (0.6)	26 (2.1)	74 (2.0)	0 (0.2)
Kentucky	41 (2.0)	56 (2.0)	3 (0.7)	24 (1.9)	75 (1.9)	1 (0.3)
Louisiana	38 (2.0)	59 (2.0)	3 (0.7)	12 (1.8)	85 (2.0)	3 (0.9)
Maine	55 (2.7)	43 (2.7)	3 (0.6)	24 (1.7)	75 (1.6)	0 (0.4)
Maryland	48 (2.2)	48 (2.3)	5 (1.0)	21 (2.0)	77 (2.0)	2 (0.7)
Massachusetts	49 (2.3)	42 (2.5)	9 (1.4)	24 (2.3)	75 (2.3)	2 (0.7)
Michigan	44 (1.9)	54 (1.8)	3 (0.7)	22 (2.6)	77 (2.6)	1 (0.4)
Minnesota	52 (1.8)	45 (1.8)	3 (0.7)	22 (2.0)	77 (2.1)	1 (0.4)
Mississippi	34 (2.0)	60 (2.1)	6 (1.0)	12 (1.5)	85 (1.6)	3 (0.8)
Missouri	46 (2.1)	51 (2.1)	3 (0.7)	21 (1.5)	77 (1.5)	1 (0.6)
Nebraska	50 (2.4)	46 (2.3)	4 (1.1)	24 (2.2)	75 (2.0)	1 (0.5)
New Hampshire	52 (2.5)	43 (2.4)	5 (0.8)	23 (2.3)	75 (2.3)	2 (0.8)
New Jersey	51 (2.1)	43 (2.3)	6 (1.2)	30 (2.0)	69 (2.1)	1 (0.5)
New Mexico	41 (2.7)	57 (2.8)	2 (0.7)	15 (2.0)	84 (2.1)	2 (0.6)
New York	45 (2.3)	50 (2.6)	6 (1.0)	17 (1.7)	80 (1.9)	3 (0.7)
North Carolina	36 (1.8)	60 (1.8)	4 (0.8)	18 (1.8)	79 (1.8)	3 (0.8)
North Dakota	53 (2.2)	44 (2.1)	3 (0.8)	25 (2.1)	73 (2.1)	1 (0.5)
Ohio	42 (2.2)	54 (2.1)	4 (0.8)	20 (2.0)	78 (2.1)	1 (0.5)
Oklahoma	47 (2.3)	49 (2.3)	4 (0.9)	20 (1.9)	79 (1.9)	1 (0.5)
Pennsylvania	53 (2.5)	45 (2.5)	3 (0.7)	21 (1.8)	78 (1.9)	1 (0.4)
Rhode Island	43 (2.2)	52 (2.3)	5 (1.0)	19 (2.2)	78 (2.4)	3 (1.0)
South Carolina	35 (2.0)	61 (2.0)	4 (0.9)	19 (1.8)	80 (1.8)	1 (0.5)
Tennessee	40 (2.3)	55 (2.2)	5 (1.0)	16 (1.6)	82 (1.6)	2 (0.6)
Texas	41 (2.1)	54 (2.3)	4 (0.8)	20 (2.0)	77 (2.2)	3 (0.7)
Utah	50 (1.9)	47 (2.1)	3 (0.7)	20 (1.4)	80 (1.5)	1 (0.3)
Virginia	45 (1.8)	51 (1.6)	4 (0.7)	22 (1.6)	77 (1.5)	2 (0.5)
West Virginia	39 (2.5)	56 (2.3)	5 (0.8)	15 (1.4)	84 (1.5)	1 (0.4)
Wisconsin	55 (2.2)	42 (2.0)	3 (0.7)	23 (1.9)	76 (2.0)	1 (0.4)
Wyoming	52 (2.4)	44 (2.4)	4 (1.0)	25 (2.1)	73 (2.1)	2 (0.9)
TERRITORY						
Guam	33 (2.5)	60 (2.7)	7 (1.3)	10 (1.5)	84 (1.9)	6 (1.4)

The standard errors of the estimated percentages appear in parentheses. It can be said with 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details). Percentages may not total 100 percent due to rounding error.

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment.

BEST COPY AVAILABLE

The following two questions also were administered only at fourth grade. For the first, classified under the content area of Numbers and Operations, 58 percent of the students supplied 7 as the correct missing digit. Many students incorrectly responded with a response of 2 and others with a response of 8. The former students were using an incorrect approach, making only unit digit analysis based on 2 x 8 resulting in a units digit of 6 in the product. The latter were most likely approaching the problem as an addition problem in the units digit, thinking that 8 + 8 results in a units digit of 6. Students who correctly answered the question had to look beyond just getting the correct units digit, to examine the related sentence 1896 ÷ 8 = 23 □.

Since developing an understanding of numerical patterns is fundamental to the introduction of algebra and functions, the second question provided fourth graders with a pattern of products, each involving a power of 2. They were asked to determine whether, if the pattern shown continued, 375 might be one of the products in the pattern. To be scored correct, students' responses had to indicate an answer of "no," and explain in some equivalent form that either 375 is not divisible by 2 (even) or that 375 falls between two numbers in the pattern (256 and 512). Nationally, only 27 percent of the fourth-grade students constructed such a response.

EXAMPLE 15: Numbers and Operations

In the multiplication problem below, write the missing number in the box.

Overall Percent Correct*
Grade 4 -- 58 (1.3)

$$\begin{array}{r} 23\boxed{7} \\ \times\ \ 8 \\ \hline 1{,}896 \end{array}$$

Did you use the calculator on this question?

(Yes) No

*The standard errors of the estimated percentages appear in parentheses.

EXAMPLE 16: Algebra and Functions

Product
$2 \times 2 = 4$
$2 \times 2 \times 2 = 8$
$2 \times 2 \times 2 \times 2 = 16$
$2 \times 2 \times 2 \times 2 \times 2 = 32$

If the pattern shown continues, could 375 be one of the products in this pattern?

Yes (No)

Explain why or why not.

Because 375 is not divisible by 2.

Did you use the calculator on this question?

(Yes) No

Overall Percent Correct*
Grade 4 -- 27 (1.5)

***The standard errors of the estimated percentages appear in parentheses.**

The subgroup and state results presented in TABLES 1.24 and 1.25 indicate that fourth graders had only a limited grasp of these situations involving multiplication. For example, the state results show that from 49 to 66 percent of the fourth graders demonstrated the ability to find the missing 7 in the multiplication problem. However, just over one-fourth were able to detect that a pattern based on multiplying 2 x 2 would involve only even numbers and articulate an answer to a problem using this information. Among the advantaged urban group, only 43 percent answered successfully, although this was significantly more than the percent correct for students attending schools in any of the three remaining types of communities. The results for the states ranged from 13 to 36 percent correct.

TABLE 1.24 National Results for Demographic Subgroups for the Regular Constructed-Response Tasks, "Missing Number in Box" and "Extend Pattern"

Grade 4

	Missing Number in Box			Extend Pattern		
	Correct	Incorrect	No Response	Correct	Incorrect	No Response
Nation	58 (1.3)	33 (1.2)	9 (0.9)	27 (1.5)	60 (1.5)	13 (0.9)
Northeast	66 (2.2)	23 (1.9)	10 (2.3)	41 (4.9)	48 (4.6)	11 (2.6)
Southeast	51 (3.5)	42 (3.1)	7 (1.4)	20 (3.6)	67 (3.7)	12 (1.3)
Central	62 (2.4)	32 (2.3)	6 (1.8)	27 (1.9)	60 (1.7)	13 (1.4)
West	56 (2.5)	32 (2.4)	12 (1.7)	23 (2.2)	61 (2.6)	16 (2.2)
White	63 (1.8)	30 (1.6)	7 (1.0)	30 (2.0)	58 (2.1)	11 (1.1)
Black	42 (3.1)	42 (3.2)	16 (2.6)	13 (2.4)	65 (3.1)	22 (3.2)
Hispanic	51 (2.5)	37 (2.8)	12 (2.6)	16 (2.9)	70 (4.2)	13 (2.4)
Male	57 (1.7)	35 (1.6)	8 (1.1)	28 (1.9)	56 (2.0)	17 (1.7)
Female	60 (1.7)	30 (1.7)	9 (1.2)	26 (1.9)	64 (2.2)	10 (1.4)
Advantaged Urban	70 (3.5)	25 (3.1)	5 (1.5)	43 (3.7)	50 (4.3)	7 (2.3)
Disadvantaged Urban	53 (4.4)	35 (4.2)	12 (2.1)	7 (2.5)	70 (3.8)	23 (4.2)
Extreme Rural	54 (5.5)	37 (4.2)	9 (2.7)	26 (5.8)	67 (5.2)	7 (2.6)
Other	58 (1.6)	33 (1.6)	9 (1.1)	26 (1.7)	59 (1.9)	14 (1.1)
Public	58 (1.5)	32 (1.4)	9 (1.0)	26 (1.6)	60 (1.7)	14 (1.1)
Catholic and Other Private	59 (2.8)	34 (2.9)	7 (1.2)	35 (3.6)	56 (4.4)	9 (2.0)

The standard errors of the estimated percentages appear in parentheses. It can be said with about 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details). Percentages may not total 100 percent due to rounding error.

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment

TABLE 1.25 | **Percentages of Correct Responses to Regular Constructed-Response Questions with Calculator Available, "Missing Number in Box" and "Extend Pattern"**

PUBLIC SCHOOLS	Grade 4 - 1992					
	Missing Number in Box			Extend Pattern		
	Correct	Incorrect	No Response	Correct	Incorrect	No Response
NATION	58 (1.5)	32 (1.4)	9 (1.0)	26 (1.6)	60 (1.7)	14 (1.1)
Northeast	67 (2.5)	23 (2.2)	10 (2.5)	42 (5.8)	48 (5.3)	10 (3.2)
Southeast	50 (3.8)	42 (3.4)	8 (1.6)	19 (4.1)	68 (4.0)	14 (1.6)
Central	62 (2.8)	31 (2.6)	7 (2.2)	24 (2.3)	61 (2.1)	14 (1.9)
West	56 (2.7)	32 (2.8)	12 (2.0)	22 (2.2)	62 (2.8)	17 (2.5)
STATES						
Alabama	52 (2.2)	40 (2.1)	7 (1.1)	20 (2.3)	64 (2.3)	16 (2.0)
Arizona	51 (2.2)	40 (2.0)	9 (1.1)	21 (1.6)	68 (2.0)	11 (1.7)
Arkansas	53 (2.3)	40 (2.0)	7 (1.1)	20 (1.9)	66 (2.3)	14 (1.7)
California	54 (2.6)	36 (2.4)	10 (1.3)	23 (2.5)	65 (2.3)	12 (1.9)
Colorado	53 (2.3)	39 (2.2)	8 (1.0)	28 (2.1)	62 (2.4)	10 (1.2)
Connecticut	60 (2.7)	31 (2.2)	9 (1.5)	32 (2.4)	61 (2.5)	8 (1.6)
Delaware	59 (1.8)	33 (1.8)	8 (1.3)	25 (2.1)	67 (2.8)	8 (1.6)
Dist. Columbia	56 (2.3)	32 (2.0)	11 (1.5)	13 (1.7)	64 (2.2)	23 (2.0)
Florida	58 (1.7)	35 (1.5)	8 (1.1)	25 (2.3)	63 (1.9)	12 (1.8)
Georgia	56 (2.0)	37 (2.1)	7 (1.1)	29 (3.0)	61 (2.9)	10 (1.7)
Hawaii	59 (2.1)	32 (1.9)	10 (1.3)	23 (2.1)	67 (2.2)	11 (1.6)
Idaho	58 (1.8)	35 (1.8)	8 (1.2)	26 (2.5)	65 (2.2)	9 (1.4)
Indiana	57 (2.5)	37 (2.5)	6 (0.9)	23 (1.9)	71 (2.4)	6 (1.3)
Iowa	60 (1.8)	36 (1.7)	4 (0.9)	35 (2.2)	58 (2.3)	7 (1.0)
Kentucky	58 (2.3)	35 (2.1)	7 (1.1)	22 (2.4)	70 (2.6)	9 (1.4)
Louisiana	58 (2.3)	34 (2.1)	7 (1.0)	15 (1.6)	72 (2.1)	12 (1.8)
Maine	61 (2.5)	34 (2.2)	5 (1.0)	35 (2.9)	56 (3.1)	10 (1.9)
Maryland	60 (2.0)	33 (1.9)	7 (1.2)	28 (2.0)	63 (2.2)	8 (1.4)
Massachusetts	61 (2.1)	28 (2.0)	11 (1.3)	35 (2.8)	56 (2.7)	9 (1.7)
Michigan	61 (2.1)	34 (2.3)	6 (0.9)	26 (3.0)	65 (2.3)	9 (1.7)
Minnesota	62 (2.1)	31 (1.9)	6 (1.1)	33 (2.6)	58 (2.9)	9 (1.4)
Mississippi	53 (2.4)	38 (2.4)	9 (1.2)	14 (1.7)	73 (2.2)	13 (1.7)
Missouri	56 (2.4)	36 (2.2)	8 (1.0)	24 (2.3)	66 (2.7)	10 (1.6)
Nebraska	60 (2.4)	35 (2.3)	6 (1.0)	33 (2.5)	61 (2.3)	6 (1.2)
New Hampshire	61 (2.7)	30 (2.1)	9 (1.5)	33 (2.9)	57 (2.9)	10 (1.7)
New Jersey	65 (2.4)	28 (1.9)	7 (1.2)	36 (2.8)	58 (3.1)	7 (1.3)
New Mexico	62 (2.7)	33 (2.5)	5 (1.1)	19 (2.8)	69 (3.7)	12 (2.3)
New York	58 (2.1)	34 (2.3)	8 (1.5)	25 (2.5)	65 (3.0)	10 (1.9)
North Carolina	57 (2.1)	36 (1.8)	8 (1.1)	22 (2.0)	66 (2.3)	12 (1.4)
North Dakota	61 (1.7)	35 (1.7)	4 (0.9)	33 (2.4)	60 (2.4)	6 (1.3)
Ohio	60 (2.0)	35 (1.9)	5 (0.9)	23 (2.2)	69 (2.4)	8 (1.1)
Oklahoma	61 (2.3)	32 (2.0)	7 (1.3)	21 (1.6)	71 (1.7)	7 (1.2)
Pennsylvania	63 (2.0)	33 (2.0)	4 (0.8)	26 (2.1)	66 (2.3)	8 (1.5)
Rhode Island	54 (2.8)	38 (2.5)	7 (1.0)	20 (2.1)	68 (2.7)	12 (1.7)
South Carolina	52 (2.0)	40 (1.8)	8 (1.1)	21 (1.9)	67 (2.3)	12 (1.7)
Tennessee	57 (2.1)	36 (1.9)	7 (1.0)	26 (2.3)	63 (2.8)	11 (1.6)
Texas	63 (1.8)	32 (1.8)	4 (0.8)	25 (2.5)	62 (2.9)	13 (1.8)
Utah	60 (1.7)	34 (1.7)	6 (1.0)	29 (2.5)	60 (2.7)	10 (1.5)
Virginia	61 (1.9)	32 (2.0)	7 (1.0)	28 (2.4)	63 (2.7)	9 (1.4)
West Virginia	53 (1.9)	40 (1.9)	7 (1.1)	24 (2.2)	62 (2.6)	14 (1.6)
Wisconsin	66 (2.1)	28 (1.9)	6 (0.9)	33 (2.4)	61 (2.3)	6 (1.2)
Wyoming	61 (2.1)	33 (1.9)	6 (0.9)	26 (2.2)	66 (2.3)	7 (1.3)
TERRITORY						
Guam	49 (2.5)	43 (2.7)	9 (1.4)	15 (2.1)	65 (2.9)	20 (2.3)

The standard errors of the estimated percentages appear in parentheses. It can be said with 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details). Percentages may not total 100 percent due to rounding error.

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment.

The following two calculator-supported tasks were presented to eighth graders as part of the national and state assessments, and to twelfth graders as part of the national assessment. The numbers and operations question asked students to calculate the number of whole packages of paper (reams) that Raymond would have to purchase in order to print 28 copies of a report containing 64 sheets of paper. Students were required to calculate the number of pages needed (1792), and round that number up to the nearest multiple of 500 (2000). For that new number, they then had to determine what multiple it was of 500 and respond that Raymond needed to purchase 4 packages of paper. If students responded with an incorrect unit, say 4 pages, they were given credit under the supposition that they meant to say 4 packages. Responses of 3, those in the range of 3.5 to 3.6, or about 4 were not scored as correct. Fifty-two percent of the eighth graders and 72 percent of the twelfth graders completed the question correctly.

The measurement task required students to determine the area of a trapezoid in square inches, given dimensions of an embedded rectangle, its area, and one other piece of necessary information. Students needed to be able to disassemble this information, using the area relationship for a rectangle to determine the altitude of the trapezoid was 4 units. Then they could proceed to use the information given to use the area formula for a trapezoid to find the area of the desired figure to be 80 square inches or use the area formula for a triangle to find the area of triangle ABE and add that area to that for the rectangle BCDE to get the 80 square inches. In either approach, students were required to manage measurement data and carry out a series of sequential calculations in order to get the desired answer. This problem proved to be quite difficult for students at both grades 8 and 12. Only 10 percent of the eighth graders and 23 percent of the twelfth graders provided a correct response to this item.

EXAMPLE 17: Numbers and Operations

Overall Percent Correct *
Grade 8 -- 52 (1.4)
Grade 12 -- 72 (1.4)

Raymond must buy enough paper to print 28 copies of a report that contains 64 sheets of paper. Paper is only available in packages of 500 sheets. How many whole packages of paper will he need to buy to do the printing?

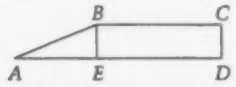
Answer: 4

Did you use the calculator on this question?

(Yes) No

EXAMPLE 18: Measurement

Overall Percent Correct *
Grade 8 -- 10 (0.9)
Grade 12 -- 23 (1.6)



The area of rectangle *BCDE* shown above is 60 square inches. If the length of *AE* is 10 inches and the length of *ED* is 15 inches, what is the area of trapezoid *ABCD*, in square inches ?

Answer: 80

Did you use the calculator on this question?

Yes (No)

*The standard errors of the estimated percentages appear in parentheses.

The subgroup results for "Raymond's Report" and "Area of a Trapezoid" for grades 8 and 12 are found in TABLE 1.26. The state results for both questions at grade 8 are found in TABLE 1.27. (State assessments were not conducted at grade 12.) In general, students had more success solving the problem of "Raymond's Report" than they did in finding the "Area of a Trapezoid." For example, across the states, from 18 to 71 percent of the eighth graders were able to find the correct answer to "Raymond's Report." In contrast, performance ranged from 1 to 16 percent correct in finding the trapezoid's area. While not set in an applied context, the "Area of a Trapezoid" problem presents a reasonable application of measurement concepts and procedures. The level of correct responses suggests that students have little grasp of how to integrate and sequence the information to arrive at a correct answer to the problem. It is interesting to note that at grade 12 twice as many private-school students as public-school students answered this question correctly -- 42 percent compared to 21 percent.

TABLE 1.26 National Results for Demographic Subgroups for the Regular Constructed-Response Tasks, "Raymond's Report" and "Area of Trapezoid"

Grade 8

	Raymond's Report			Area of Trapezoid		
	Correct	Incorrect	No Response	Correct	Incorrect	No Response
Nation	52 (1.4)	46 (1.4)	3 (0.4)	10 (0.9)	81 (1.3)	9 (0.8)
Northeast	58 (3.5)	38 (3.5)	4 (0.7)	9 (1.4)	80 (1.7)	11 (1.5)
Southeast	42 (2.9)	54 (3.1)	4 (1.1)	9 (1.9)	82 (2.8)	9 (1.3)
Central	61 (3.3)	39 (3.4)	0 (0.2)	10 (1.9)	84 (2.1)	5 (0.8)
West	48 (1.9)	49 (1.6)	3 (1.0)	10 (2.0)	79 (3.0)	11 (2.3)
White	62 (1.7)	36 (1.7)	2 (0.4)	12 (1.2)	80 (1.4)	8 (0.7)
Black	20 (3.2)	74 (3.5)	6 (1.1)	2 (1.1)	84 (3.8)	14 (3.3)
Hispanic	30 (3.7)	65 (3.8)	5 (1.4)	3 (1.3)	88 (2.3)	9 (1.9)
Male	51 (2.3)	46 (2.4)	3 (0.6)	10 (1.4)	79 (2.0)	12 (1.4)
Female	52 (1.9)	45 (1.9)	2 (0.5)	9 (1.2)	84 (1.5)	7 (0.9)
Advantaged Urban	66 (5.3)	34 (5.3)	0 (0.0)	19 (3.8)	70 (4.0)	11 (4.9)
Disadvantaged Urban	25 (4.5)	68 (4.8)	6 (1.7)	4 (1.3)	82 (2.8)	14 (3.1)
Extreme Rural	55 (8.2)	45 (8.1)	0 (0.4)	9 (3.3)	83 (4.8)	8 (2.7)
Other	52 (1.6)	44 (1.6)	3 (0.6)	9 (1.2)	83 (1.5)	9 (0.7)
Public	50 (1.5)	47 (1.5)	3 (0.5)	9 (1.0)	82 (1.4)	9 (0.9)
Catholic and Other Private	62 (3.2)	35 (3.2)	3 (0.6)	12 (2.0)	78 (2.5)	9 (1.6)

Grade 12

	Raymond's Report			Area of Trapezoid		
	Correct	Incorrect	No Response	Correct	Incorrect	No Response
Nation	72 (1.4)	25 (1.4)	2 (0.5)	23 (1.6)	67 (1.6)	10 (0.9)
Northeast	75 (2.2)	23 (2.3)	2 (0.9)	26 (2.8)	64 (3.4)	10 (1.6)
Southeast	68 (2.8)	28 (2.7)	3 (1.1)	16 (2.3)	77 (2.2)	7 (1.5)
Central	78 (1.9)	21 (2.2)	2 (0.9)	26 (3.2)	64 (2.1)	10 (2.4)
West	69 (3.5)	29 (3.7)	2 (0.6)	24 (4.2)	65 (4.3)	11 (1.4)
White	78 (1.3)	20 (1.3)	1 (0.3)	27 (2.1)	63 (1.8)	10 (1.1)
Black	51 (4.2)	44 (4.7)	5 (1.5)	8 (1.9)	81 (3.0)	12 (2.6)
Hispanic	62 (5.9)	34 (6.2)	4 (2.4)	14 (2.3)	80 (2.5)	6 (1.8)
Male	74 (2.0)	23 (2.0)	4 (0.9)	24 (1.7)	65 (1.8)	10 (1.2)
Female	71 (1.9)	28 (1.9)	1 (0.3)	22 (2.4)	68 (2.2)	9 (1.3)
Advantaged Urban	79 (3.6)	21 (3.7)	0 (0.2)	46 (4.8)	50 (4.6)	4 (1.4)
Disadvantaged Urban	62 (3.4)	33 (3.8)	5 (1.4)	10 (2.1)	79 (2.7)	11 (2.3)
Extreme Rural	73 (3.8)	25 (3.5)	2 (1.0)	25 (4.0)	68 (2.7)	8 (2.4)
Other	73 (1.5)	25 (1.6)	2 (0.6)	21 (1.6)	68 (1.9)	11 (1.0)
Public	72 (1.5)	26 (1.5)	2 (0.5)	21 (1.9)	70 (1.8)	10 (0.9)
Catholic and Other Private	79 (2.4)	20 (2.4)	2 (0.8)	42 (3.4)	49 (3.5)	9 (1.8)

The standard errors of the estimated percentages appear in parentheses. It can be said with about 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details). When the proportion of students is either 0 percent of 100 percent, the standard error is inestimable. However, percentages 99.5 percent and greater were rounded to 100 percent and percentages 0.5 percent or less ere rounded to 0 percent. Percentages may not total 100 percent due to rounding error.

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment

TABLE 1.27 | **Percentages of Correct Responses to Regular Constructed-Response Questions with Calculator Available, "Raymond's Report" and "Area Trapezoid"**

PUBLIC SCHOOLS	Grade 8 - 1992					
	Raymond's Report			Area of Trapezoid		
	Correct	Incorrect	No Response	Correct	Incorrect	No Response
NATION	50 (1.5)	47 (1.5)	3 (0.5)	9 (1.0)	82 (1.4)	9 (0.9)
Northeast	58 (4.1)	39 (4.1)	3 (0.8)	7 (1.9)	81 (2.2)	12 (1.7)
Southeast	38 (2.5)	58 (2.9)	4 (1.2)	9 (2.1)	82 (3.0)	9 (1.4)
Central	60 (3.7)	39 (3.7)	0 (0.2)	10 (2.3)	85 (2.3)	5 (1.0)
West	47 (2.0)	49 (1.7)	4 (1.1)	9 (2.0)	79 (3.2)	11 (2.5)
STATES						
Alabama	41 (2.3)	56 (2.3)	3 (0.7)	4 (0.7)	89 (1.5)	7 (1.3)
Arizona	53 (2.2)	44 (2.2)	3 (0.7)	9 (1.3)	82 (1.8)	9 (1.2)
Arkansas	43 (1.9)	54 (2.0)	3 (0.8)	5 (0.8)	89 (1.2)	6 (0.9)
California	48 (2.3)	47 (2.2)	5 (1.0)	10 (1.3)	76 (1.5)	14 (1.4)
Colorado	56 (1.9)	41 (1.9)	3 (0.8)	9 (1.4)	81 (1.5)	10 (1.2)
Connecticut	58 (2.2)	40 (2.3)	2 (0.6)	12 (1.4)	81 (1.8)	7 (1.1)
Delaware	54 (2.7)	43 (2.3)	2 (0.9)	6 (1.3)	85 (1.7)	9 (1.3)
Dist. Columbia	30 (2.8)	62 (2.7)	8 (1.2)	3 (1.1)	87 (2.2)	10 (1.9)
Florida	52 (2.1)	44 (2.1)	4 (1.0)	6 (0.9)	85 (1.6)	9 (1.4)
Georgia	45 (2.3)	52 (2.6)	3 (0.9)	5 (1.0)	90 (1.3)	4 (0.8)
Hawaii	48 (2.1)	48 (2.1)	4 (0.8)	8 (1.2)	79 (1.9)	13 (1.5)
Idaho	58 (1.7)	39 (1.6)	2 (0.6)	13 (1.6)	79 (1.8)	8 (1.0)
Indiana	55 (2.0)	43 (2.0)	2 (0.7)	9 (1.3)	84 (1.6)	7 (1.1)
Iowa	71 (2.2)	28 (2.3)	2 (0.5)	13 (1.5)	82 (1.5)	5 (1.0)
Kentucky	54 (2.3)	43 (2.2)	3 (0.7)	7 (1.1)	87 (1.4)	6 (0.9)
Louisiana	42 (2.5)	54 (2.4)	5 (1.1)	4 (1.1)	90 (1.6)	7 (1.5)
Maine	67 (2.3)	32 (2.3)	2 (0.4)	12 (1.5)	80 (1.9)	8 (1.3)
Maryland	54 (2.4)	42 (2.2)	4 (1.0)	9 (1.5)	81 (2.0)	10 (1.4)
Massachusetts	59 (2.6)	39 (2.5)	3 (0.8)	9 (1.1)	82 (1.4)	8 (1.1)
Michigan	55 (2.1)	43 (2.2)	2 (0.7)	10 (1.4)	81 (1.6)	8 (1.2)
Minnesota	66 (1.7)	32 (1.7)	2 (0.6)	15 (1.7)	78 (1.9)	7 (1.0)
Mississippi	35 (2.1)	61 (2.2)	4 (0.9)	3 (0.7)	89 (1.1)	8 (1.0)
Missouri	53 (1.9)	46 (1.9)	1 (0.5)	8 (1.2)	85 (1.6)	7 (1.2)
Nebraska	58 (2.3)	41 (2.3)	2 (0.6)	12 (1.5)	83 (1.8)	4 (0.8)
New Hampshire	63 (2.1)	36 (2.1)	1 (0.5)	13 (1.5)	78 (1.9)	8 (1.3)
New Jersey	59 (3.2)	37 (3.1)	3 (1.0)	9 (1.7)	84 (2.3)	7 (1.2)
New Mexico	46 (2.2)	52 (2.2)	2 (0.6)	7 (1.1)	85 (1.6)	7 (1.2)
New York	55 (2.4)	43 (2.4)	2 (0.8)	12 (1.7)	84 (1.8)	4 (0.9)
North Carolina	48 (1.9)	50 (1.8)	2 (0.6)	5 (0.9)	90 (1.2)	6 (0.8)
North Dakota	67 (2.0)	32 (2.0)	1 (0.5)	16 (1.8)	79 (1.9)	5 (1.0)
Ohio	58 (2.3)	41 (2.2)	1 (0.4)	6 (1.0)	87 (1.4)	6 (1.0)
Oklahoma	59 (2.5)	39 (2.5)	2 (0.5)	8 (1.2)	85 (1.4)	6 (1.2)
Pennsylvania	56 (2.5)	41 (2.6)	2 (0.7)	10 (1.3)	84 (1.9)	7 (1.4)
Rhode Island	54 (2.2)	45 (2.0)	2 (0.6)	6 (1.0)	88 (1.4)	6 (0.8)
South Carolina	44 (2.5)	54 (2.4)	3 (0.7)	8 (1.1)	86 (1.4)	6 (0.9)
Tennessee	44 (2.4)	52 (2.4)	4 (0.7)	5 (1.0)	88 (1.4)	7 (1.0)
Texas	51 (2.3)	45 (2.2)	4 (0.7)	9 (1.2)	83 (1.6)	8 (1.2)
Utah	61 (1.8)	38 (1.8)	1 (0.4)	10 (1.3)	83 (1.4)	7 (1.1)
Virginia	57 (2.1)	42 (2.0)	2 (0.6)	8 (1.1)	86 (1.3)	6 (1.1)
West Virginia	52 (2.2)	46 (2.1)	2 (0.5)	5 (0.8)	88 (1.3)	8 (1.1)
Wisconsin	65 (1.7)	34 (1.5)	1 (0.4)	11 (1.3)	83 (1.3)	6 (0.9)
Wyoming	61 (2.2)	37 (2.3)	2 (0.8)	9 (1.1)	86 (1.3)	6 (0.8)
TERRITORIES						
Guam	29 (2.8)	64 (2.7)	7 (1.3)	4 (1.0)	87 (1.8)	9 (1.7)
Virgin Islands	18 (1.8)	68 (2.2)	14 (2.0)	1 (0.5)	82 (2.1)	17 (2.2)

The standard errors of the estimated percentages appear in parentheses. It can be said with 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details). Percentages may not total 100 percent due to rounding error.

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment.

The remaining five calculator-aided, constructed-response questions presented in this report were administered only at grade 12. In "Video Rental Costs," twelfth graders were presented with a comparison shopping situation involving two video rental stores. The problem was complicated by the presence of a bonus rental feature which provided free tapes under differing rental conditions. While it was possible for students to get the rental cost for one of the two stores correct and miss the other, credit for the item was only given to students getting both of the rental costs correct.

The question asking students to find the area of a parallelogram was classified in the Measurement content area. Students were given sufficient information to find the solution, but any approach to the problem required that students make an application of the Pythagorean theorem to find the length of the missing segment on the base of the parallelogram. Once this segment was found and added to the 7 units shown for the other portion of the base, the student only needed to multiply the sum by the altitude 7 to obtain the rounded area of 188 or 189 square units. A frequent incorrect answer was 91, from 7 x 13, where students neglected to find the other portion of the base for the parallelogram.

As can be seen from the results in TABLE 1.28, these two multi-step problems gave twelfth graders considerable difficulty. Only 5 percent answered the "Video Rental Costs" question correctly, and only 8 percent correctly determined the area of the parallelogram.

EXAMPLE 19: Numbers and Operations

Overall Percent Correct *
Grade 12 -- 5 (0.7)

Video Store A	Video Store B
$2.65 per tape for one night $1.50 charge for each additional night Every 10th tape **free** for one night	$3.00 per tape for 2 nights 1 credit if tape returned after one night Every 10 credits = one **free** rental

The Peterson family rents 30 videotapes yearly, of which 23 are rented for one night only and 7 are rented over a period of two nights. Given the rental fee structures shown above, fill in the chart below with the total yearly cost for the Petersons at each store. (Note: The 30 tapes include the free tapes earned.)

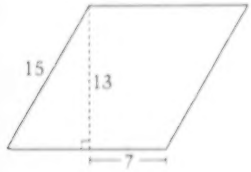
Store	Total Cost
A	$82.05
B	$84.00

Did you use the calculator on this question?

(Yes) No

EXAMPLE 20: Measurement

Overall Percent Correct *
Grade 12 -- 8 (1.0)



To the nearest whole number, what is the area of the parallelogram above?

Answer: 188

Did you use the calculator on this question?

(Yes) No

*The standard errors of the estimated percentages appear in parentheses.

TABLE 1.28 National Results for Demographic Subgroups for the Regular Constructed-Response Tasks, "Video Rental Costs" and "Area of Parallelogram"

Grade 12

	Video Rental Costs			Area of Parallelogram		
	Correct	Incorrect	No Response	Correct	Incorrect	No Response
Nation	5 (0.7)	92 (0.9)	3 (0.5)	8 (1.0)	80 (1.4)	12 (1.1)
Northeast	8 (2.5)	90 (3.1)	2 (0.7)	9 (2.7)	78 (4.4)	13 (2.3)
Southeast	6 (1.1)	92 (1.6)	3 (1.1)	4 (0.6)	86 (2.0)	10 (2.0)
Central	4 (1.0)	96 (1.1)	1 (0.3)	10 (2.0)	79 (2.1)	11 (2.1)
West	4 (1.4)	90 (1.9)	6 (1.3)	8 (2.1)	77 (2.3)	14 (2.2)
White	6 (0.9)	92 (1.0)	2 (0.5)	9 (1.3)	80 (1.7)	11 (1.1)
Black	2 (1.2)	94 (2.2)	4 (1.5)	1 (0.5)	85 (2.6)	15 (2.6)
Hispanic	0 (0.0)	92 (4.5)	8 (4.5)	2 (1.4)	77 (4.3)	21 (4.5)
Male	6 (0.9)	90 (1.4)	4 (1.0)	8 (1.4)	81 (1.7)	11 (1.2)
Female	5 (1.1)	94 (1.5)	2 (0.6)	7 (1.2)	79 (2.1)	14 (1.8)
Advantaged Urban	8 (3.4)	89 (4.0)	3 (1.7)	14 (2.9)	74 (3.4)	11 (1.8)
Disadvantaged Urban	6 (2.4)	88 (4.1)	7 (2.1)	1 (0.5)	85 (2.7)	14 (2.5)
Extreme Rural	5 (1.3)	93 (1.9)	2 (1.0)	6 (2.2)	82 (4.2)	12 (4.4)
Other	5 (0.9)	93 (1.1)	2 (0.7)	8 (1.5)	80 (1.9)	12 (1.4)
Public	5 (0.7)	92 (0.9)	3 (0.5)	7 (0.8)	81 (1.2)	13 (1.2)
Catholic and Other Private	7 (2.2)	91 (2.3)	2 (1.0)	15 (4.2)	75 (5.3)	10 (1.7)

The standard errors of the estimated percentages appear in parentheses. It can be said with about 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details). When the proportion of students is either 0 percent or 100 percent, the standard error is inestimable. However, percentages 99.5 percent and greater were rounded to 100 percent and percentages or less were rounded to 0 percent. Percentages may total 100 percent due to rounding error.

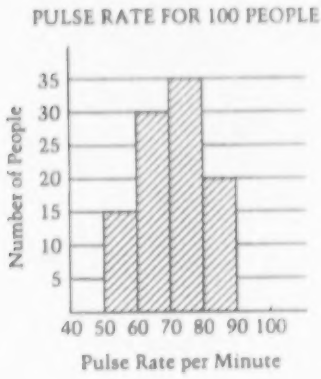
SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment

Of the next two tasks, one is a data analysis, statistics, and probability question requiring students to find the average pulse rate for a group of individuals. Information on these 100 individuals was grouped in the form of a histogram having pulses per minute rates calibrated in terms of 10 pulses. This required students to consider the pulses in a given category to be clustered at the average value (or midpoint) of each interval. Students were given a hint to consider this approach to dealing with the data as part of the problem statement.

The algebra and functions question required students to understand the meaning and application of the $f(x)$ notation, and make the substitution of 3.5 for x in the function rule for f. This approach would result in the calculation $4(3.5)^2 - 7(3.5) + 5.7$, or 30.2.

EXAMPLE 21: Data Analysis, Statistics, and Probability

Overall Percent Correct*
Grade 12 – 9 (1.0)



The pulse rate for a group of 100 people is shown in the graph above. What is the average pulse rate per minute for these 100 people? (Note: Use the midpoint of each interval to represent the pulse rate for the entire interval. For example, 55 would be used for the pulse rate of the 15 people in the 50-60 group.)

Answer: 71

Did you use the calculator on this question?

(Yes) No

EXAMPLE 22: Algebra and Functions

Overall Percent Correct*
Grade 12 – 39 (1.6)

If $f(x) = 4x^2 - 7x + 5.7$, what is the value of $f(3.5)$?

Answer: 30.2

Did you use the calculator on this question?

(Yes) No

*The standard errors of the estimated percentages appear in parentheses.

As shown in TABLE 1.29, students had considerable difficulty with the "Graph of Pulse Rates." Only 9 percent determined the average pulse rate from the histogram. More twelfth graders were familiar with what was needed to complete the functional notation task, as 39 percent of the students received credit for this item.

TABLE 1.29 National Results for Demographic Subgroups for the Regular Constructed-Response Tasks, "Graph Pulse Rates" and "F(3.5)"

Grade 12

	Graph Pulse Rates			F(3.5)		
	Correct	Incorrect	No Response	Correct	Incorrect	No Response
Nation	9 (1.0)	77 (1.3)	14 (1.0)	39 (1.6)	40 (1.5)	22 (1.4)
Northeast	10 (2.1)	74 (2.5)	16 (1.8)	40 (2.5)	35 (2.2)	24 (2.9)
Southeast	5 (1.1)	79 (2.4)	17 (2.4)	31 (2.6)	46 (3.4)	23 (3.2)
Central	10 (1.7)	80 (1.7)	10 (1.4)	45 (4.0)	37 (3.2)	18 (2.4)
West	9 (2.5)	77 (2.7)	14 (1.8)	39 (2.6)	40 (3.0)	21 (2.4)
White	11 (1.3)	75 (1.5)	14 (1.1)	43 (1.8)	36 (1.3)	21 (1.6)
Black	0 (0.0)	86 (2.9)	14 (2.9)	23 (4.1)	55 (4.9)	22 (3.5)
Hispanic	1 (0.6)	83 (3.3)	16 (3.2)	25 (5.7)	44 (9.0)	31 (5.9)
Male	11 (1.6)	75 (1.8)	14 (1.3)	37 (2.1)	40 (2.1)	23 (1.9)
Female	6 (1.0)	80 (1.9)	14 (1.7)	41 (2.0)	39 (1.7)	20 (1.9)
Advantaged Urban	19 (4.4)	69 (4.7)	12 (3.6)	58 (5.3)	26 (3.0)	16 (4.0)
Disadvantaged Urban	1 (0.8)	84 (3.1)	15 (2.8)	25 (3.7)	50 (4.1)	25 (3.4)
Extreme Rural	8 (2.3)	77 (2.4)	16 (3.2)	28 (6.0)	51 (5.1)	21 (4.9)
Other	8 (1.1)	78 (1.7)	14 (1.4)	40 (1.7)	38 (1.8)	22 (1.5)
Public	8 (1.1)	78 (1.4)	14 (1.2)	37 (1.7)	40 (1.7)	22 (1.6)
Catholic and Other Private	15 (2.7)	74 (3.1)	10 (1.7)	50 (3.0)	33 (2.1)	17 (1.9)

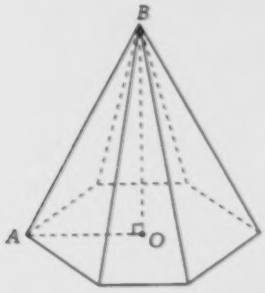
The standard errors of the estimated percentages appear in parentheses. It can be said with about 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details). When the proportion of students is either 0 percent of 100 percent, the standard error is inestimable. However percentages 99.5 percent and greater were rounded to 100 percent and percentages 0.5 percent or less were rounded to 0 percent. Percentages may not total 100 percent due to rounding error.

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment

The final calculator-assisted constructed-response question given to grade 12 students was an algebra and functions item that required students to make use of a number of relationships drawn from measurement, geometry, algebra, and trigonometry. To correctly answer the item, the student had to recognize that triangle OAB is a right triangle and that since the regular hexagon is composed

of 6 equilateral triangles, the length of segment AO is found to be 12. Then the measure of angle OAB to the nearest whole number is the rounded value of Tan^{-1} (15/12) or 51°. Some students gave the answer in terms of radian measure 0.8960..., which was incorrect, as the problem required degree measure. The results presented in TABLE 1.30 reveal that, for the nation, 7 percent of the twelfth graders correctly solved the problem. Even for those students attending schools in advantaged urban communities, only 14 percent provided correct responses. Catholic- and other private-school students performed significantly better than public-school students, 13 compared to 6 percent correct, respectively.

EXAMPLE 23: Algebra and Functions



Overall Percent Correct*
Grade 12 – 7 (0.5)

The base of the pyramid shown above is a regular hexagon with side of length 12. If point *O* is the center of the base and the length of *OB* is 15, what is the degree measure of angle *OAB* to the nearest whole number?

Answer: 51°

Did you use the calculator on this question?

(Yes) No

*The standard errors of the estimated percentages appear in parentheses.

TABLE 1.30 **National Results for Demographic Subgroups for the Regular Constructed-Response Task, "Trigonometry"**

Grade 12

	Correct	Incorrect	No Response
Nation	7 (0.5)	84 (1.0)	10 (0.8)
Northeast	8 (1.3)	80 (2.0)	12 (1.7)
Southeast	3 (0.8)	89 (2.2)	8 (1.7)
Central	10 (1.2)	82 (1.8)	9 (1.5)
West	5 (1.2)	86 (1.8)	10 (1.7)
White	8 (0.7)	83 (1.2)	10 (1.0)
Black	1 (0.9)	90 (2.5)	9 (2.4)
Hispanic	2 (1.2)	90 (4.0)	8 (3.3)
Male	8 (1.0)	82 (1.3)	10 (1.1)
Female	6 (1.0)	85 (1.5)	9 (1.1)
Advantaged Urban	14 (2.3)	76 (3.2)	10 (2.7)
Disadvantaged Urban	2 (1.0)	86 (3.1)	12 (2.6)
Extreme Rural	3 (1.1)	88 (3.7)	9 (3.8)
Other	7 (0.8)	84 (1.3)	9 (1.0)
Public	6 (0.7)	85 (1.1)	10 (0.9)
Catholic and Other Private	13 (2.2)	79 (2.4)	8 (1.2)

The standard errors of the estimated percentages appear in parentheses. It can be said with about 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details). Percentages may total 100 percent due to rounding error.

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment

Summary of National and State Results for Regular Constructed-Response Items

In general, student attempts to construct their answers to the questions presented in this chapter suggested honest efforts to comply with the requirements set by the problem situations. However, the information provided by students added up to generally low performance levels. Students had difficulty with questions in all mathematics content areas. Those requiring "hands-on" measurement skills using a ruler or protractor may have implications for applications in various technological and daily-life settings. Because students were supplied with calculators to use in answering some of the questions, these tasks required slightly more of students in the way of both computation and the number of steps involved. Students had particular difficulty with these questions. Even though students had access to the calculator in working on these problems,

they did not demonstrate an ability to conceptualize or sequence the information required by the problem-solving situation.

In general, performance for subpopulations followed patterns found in other NAEP mathematics results, with greater percentages of White students than Black and Hispanic students, as well as higher percentages of advantaged urban than disadvantaged urban students, providing correct responses to these questions. Gender and regional differences were less consistent from question to question as were results for public- versus private-school students (see Chapter Three for summary results).

At grades 4 and 8, the results for participating states and territories tended to mirror the low levels of national performance, even though there was considerable variation across the jurisdictions.

CHAPTER TWO

Extended Responses to Explain Mathematical Reasoning

Introduction

The constructed-response tasks presented in this chapter represent a step toward the mathematics assessments envisioned for the future -- assessments that reflect instructional goals to actively promote student learning.[11] The tasks not only require students to construct their own responses instead of choosing a single answer, but, in contrast to the questions described in Chapter One, they also provide students an opportunity to express their mathematical ideas and demonstrate the depth of their understanding of a problem. These types of tasks are intended as progress toward addressing *The NCTM Standards*, which emphasize that students demonstrate their problem-solving and reasoning abilities and learn to communicate effectively about the mathematical power they possess. The intention is for NAEP to continue making strides in improving the problem-solving tasks given to students with each successive assessment, building on experience with performance-oriented assessment approaches to incorporate improved procedures in the future.[12]

Overview of the Tasks and Scoring Guides

Particularly in a large-scale assessment situation, it is a great challenge to develop tasks that exemplify *The NCTM Standards*, but remain within the reach of most students. Further, issues arise regarding the operational aspects of providing the ancillary materials that can be used with such problem situations, the time such tasks take, and the reliability of the scoring. A tightly structured question can be scored more easily, but in some instances yields less interesting information about students' mathematical understanding. On the other hand,

---

[11] Mathematical Sciences Education Board, *For Good Measure* (Washington, DC: National Academy Press, 1991).

[12] *1994 National Assessment of Educational Progress: Mathematics Assessment Framework* (Washington, DC: National Assessment Governing Board, U.S. Government Printing Office, 1993).

more broadly structured questions require developing guides that take into account the strategies used to solve the problem, the implementation of those strategies, and the interaction between the two. Each task needs to be considered separately in developing scoring rubrics or guides, because both the strategies and implementation may differ significantly across problem situations and content. As might be anticipated, the first round of extended problem-solving tasks developed by NAEP represents a range in the structure provided, the settings and stimulus materials encompassed by the problem situations, and in the content covered.

At each of grades 4, 8, and 12, three examples of extended tasks from the 1992 assessment are presented, along with the associated scoring rubrics and examples of student work. Each of these extended constructed-response questions required students to demonstrate their level of mathematical understanding within a given context that varied from question to question -- both in terms of the approach to the problem and its content domain. Students were asked to think carefully about the question before writing a complete answer that demonstrated their understanding of the problem. They were asked to show all the work that led to their solution, or to provide an explanation of their reasoning. In formulating their responses to these extended questions, students were told that they could use drawings, words, and numbers in their explanations and that it was important that their solution be clear enough so that another person could read it and understand their thinking.

To provide for some consistency in approach for developing the evaluation criteria across the extended-response tasks, five generic levels of performance were defined as shown in FIGURE 2.1. Provision also was made for categorizing blank papers that provided no response to the question. These guiding principles were used in developing the tailored scoring guides for each specific task.

FIGURE 2.1 NAEP Mathematics Scoring Category Classifications for Extended Constructed-Response Questions

0 =	*No Response*	There is no response
1 =	*Incorrect Response*	The work is completely incorrect or irrelevant. Or the response states, "I don't know."
2 =	*Minimal*	The response demonstrates a minimal understanding of the problem posed but does not suggest a reasonable approach. Although there may or may not be some correct mathematical work, the response is incomplete, contains major mathematical errors, or reveals serious flaws in reasoning. Examples are absent.
3 =	*Partial*	The response contains evidence of a conceptual understanding of the problem in that a reasonable approach is indicated. However, on the whole, the response is not well developed. Although there may be serious mathematical errors or flaws in the reasoning, the response does contain some correct mathematics. Examples provided are inappropriate.
4 =	*Satisfactory*	The response demonstrates a clear understanding of the problem and provides an acceptable approach. The response also is generally well developed and coherent but contains minor weaknesses in the development. Examples provided are not fully developed.
5 =	*Extended*	The response demonstrates a complete understanding of the problem, is correct, and the methods of solution are appropriate and fully developed. Responses scored 5 are logically sound, clearly written, and do not contain any significant mathematical errors. Examples are well chosen and fully developed.

NOTE: In the partial-credit scaling used to summarize student achievement across questions for other analyses and reports, neither "blank" nor incorrect responses are given credit, and minimal through extended responses are assigned values 1-4. However, for the purposes of this report, it is interesting to distinguish between students who omitted questions and those who at least attempted to respond.

Organization for this Chapter

The national and state results for each of the nine questions are presented in the following format:

- **The Task** -- The problem situation given to students is presented.
- **Possible Solution** -- An overview of the "ingredients" of a successful approach is provided together with one possible example of a successful solution.

BEST COPY AVAILABLE

- **National Results, Scoring Guide, and Sample Responses** -- The national results are presented category by category, together with the description of the scoring guide for that category, and an annotated example student response illustrative of the types of answers students gave.

- **National Results for Demographic Subgroups** -- The results for subpopulations of students are presented as defined by region, race/ethnicity, gender, type of community, and type of school (students attending public schools as compared to those attending private schools, including Catholic and other types of private schools).

- **State Results** -- The state-by-state results for grades 4 and 8 are shown for each category of performance, including the percentage of satisfactory or better responses (categories 4 and 5). It should be noted that for comparisons between the nation or the regions and the participating states, the national and regional data provided in the state tables should be used. In contrast to the state assessments, which only included students attending public schools, the national assessments included students attending both public and private schools. Thus, for comparison purposes, the national and regional results presented together with the state results are based only on public-school students, while those presented earlier are based on students attending both public and private schools.

- **Performance Highlights** -- The national and state results are discussed.

Extended-Response Questions: Grade 4

The fourth-grade students who participated in the 1992 assessment were given five extended-response tasks, three of which follow.

The "Pizza Comparison" question assesses how well students are making the transition from whole number reasoning into using concepts associated with fractions. The intent of the problem is to measure students' ability to communicate that a fraction must be interpreted in terms of the relative size of the object. It taps their understanding of the concept of fraction with initial development of the idea of proportional reasoning. The real-life setting for this numbers and operations question pertains to comparing pieces of pizza.

In "Laura's Calculator Correction," students were provided with a four-function calculator (TI-108) and asked to apply their understanding of place value to explain two ways for correcting a mistaken entry -- 8375 instead of 8275 -- without clearing the calculator.

The "Graphs of Pockets" question asks students to read, interpret, and select one of three pictographs to represent a particular situation -- the number of pockets for a class of 20 students. In this data analysis task, students are asked to explain why they selected a particular graph and their reasons for rejecting the other two.

Grade 4 Question: Pizza

The Task

Think carefully about the following question. Write a complete answer. You may use drawings, words, and numbers to explain your answer. Be sure to show all of your work.

José ate $\frac{1}{2}$ of a pizza.

Ella ate $\frac{1}{2}$ of another pizza.

José said that he ate more pizza than Ella, but Ella said they both ate the same amount. Use words and pictures to show that José could be right.

Possible Solution

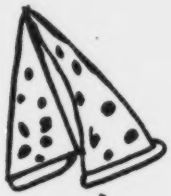
Jose would be right if the size of his pizza was larger than the size of Ella's pizza. More generally, students are expected to communicate by pictures and/or words that half of a larger quantity is more than half of a smaller quantity.

Students with only a naive understanding of the meaning of "1/2" in the context of the given task are likely to indicate "1/2 = 1/2" because they do not realize the potential for the two quantities being compared, the pizzas, to be different in size. Students with a higher level of comprehension can show some evidence that size is an important factor but are unable to convey how the comparison of the two pizzas is related to their relative sizes. Students with the highest level of understanding of the meaning of "1/2" in the context of the given problem can demonstrate responses that, at least informally, demonstrate what the fraction 1/2 means in terms of relative sizes of pizzas.

National Results, Scoring Guide, and Sample Responses

National Percent for Each Category*	Rating	Rating and Performance Category
7 (0.8)	0	No Response
49 (1.7)	1	Incorrect -- The work is completely incorrect or irrelevant, or the response states, "I don't know."

This INCORRECT response does not involve the concept of one-half of a whole pizza.



José ate 1/2

18 (1.1)	2	Minimal -- Student responds that "1/2 is always 1/2" indicating an awareness of fractional parts. Other responses may include only references to number of pizzas or to toppings.

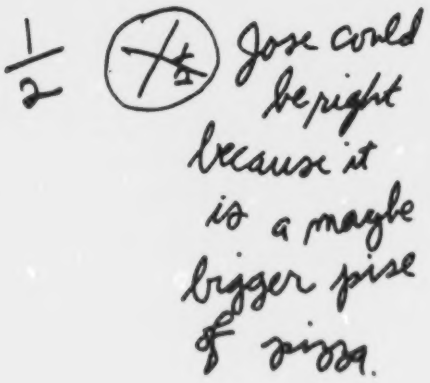
This MINIMAL response indicates an understanding of the concept of 1/2 as a fractional part of a whole, but states 1/2 is always equal to 1/2.

Jose ate his 1/2 and Ella ate her 1/2 they both had 1/2 and they both ate the same amount.

* The standard errors of the estimated percentages appear in parentheses.

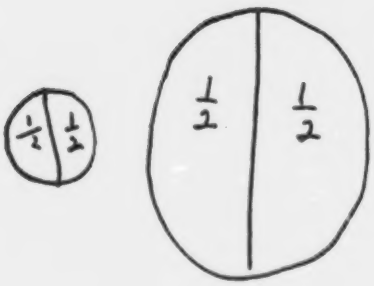
National Percent for Each Category	Rating and Performance Category	
2 (0.5)	3	**Partial -- Student makes statements such as "Jose's pizza has bigger pieces" that begin to demonstrate an awareness of the idea of relative size.**

This PARTIAL response does give an indication that Jose's pizza may be larger.



8 (0.8)	4	**Satisfactory -- Student displays responses that connect figuraly the relationship between the difference in the relative size of Jose's and Ella's pizzas but are not clear in explaining that relationship.**

This SATISFACTORY response uses diagrams to clearly show two different-sized pizzas and to illustrate that the respective halves of those pizzas are not the same size.



National Percent for Each Category

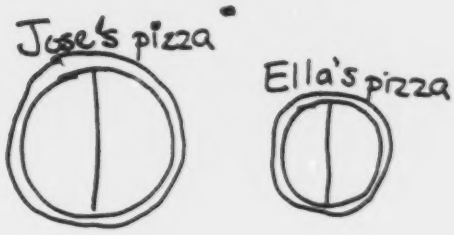
16 (1.2)

Rating and Performance Category

5 **Extended -- Student explains and/or demonstrates a clear understanding of fractional part and relative size.**

This strong EXTENDED response provides drawings of two different-sized pizzas, each divided into halves and labelled appropriately. The student also has written a clear and accurate description of the situation.

José could be right because his pizza could be bigger than Ella's.



Performance Highlights: Pizza Comparison

Nationally, the majority of fourth graders appeared to be unfamiliar with either the format of such a mathematics task (7 percent provided no response) or the concepts underlying the question. As shown in TABLE 2.1, nearly half of the students (49 percent) did not use or include mathematical concepts in communicating their responses. Of those students who provided irrelevant or incorrect answers, most tended either to draw pizzas or parts of them or simply to reiterate the premise in the problem, showing that either José, Ella, or both ate half a pizza.

TABLE 2.1 National Results for Demographic Subgroups for the Extended-Response Question, "Pizza Comparison"

Grade 4

	No Response	Incorrect	Minimal	Partial	Satisfactory	Extended	Satisfactory or Better
Nation	7 (0.8)	49 (1.7)	18 (1.1)	2 (0.5)	8 (0.8)	16 (1.2)	23 (1.3)
Northeast	7 (2.1)	42 (4.6)	19 (3.2)	3 (1.2)	8 (2.3)	21 (3.5)	29 (4.3)
Southeast	6 (1.1)	55 (4.2)	17 (2.1)	2 (0.6)	6 (1.2)	14 (2.0)	20 (2.4)
Central	7 (1.7)	49 (2.8)	18 (1.7)	3 (1.3)	8 (2.0)	15 (2.6)	23 (2.2)
West	8 (1.6)	50 (3.3)	18 (2.0)	2 (0.8)	9 (1.1)	14 (2.0)	23 (2.1)
White	6 (0.9)	44 (2.1)	20 (1.3)	2 (0.6)	9 (1.2)	19 (1.5)	28 (1.7)
Black	11 (2.7)	65 (3.5)	13 (2.3)	2 (1.0)	6 (1.9)	3 (1.3)	9 (2.1)
Hispanic	7 (1.9)	64 (3.6)	16 (3.1)	1 (0.5)	5 (1.7)	8 (2.3)	12 (2.8)
Male	9 (1.2)	48 (2.4)	15 (1.3)	2 (0.5)	8 (1.3)	17 (2.0)	26 (2.0)
Female	5 (0.9)	51 (2.2)	20 (1.7)	3 (0.8)	7 (1.0)	14 (1.2)	21 (1.4)
Advantaged Urban	3 (1.3)	41 (4.1)	16 (2.7)	4 (1.6)	10 (2.6)	26 (4.0)	35 (3.6)
Disadvantaged Urban	8 (2.2)	68 (5.3)	10 (2.3)	0 (0.3)	10 (3.5)	4 (1.6)	14 (3.5)
Extreme Rural	7 (2.3)	55 (4.6)	17 (2.9)	2 (1.2)	4 (1.5)	14 (3.5)	18 (3.7)
Other	8 (1.1)	47 (2.2)	20 (1.3)	2 (0.7)	8 (0.9)	16 (1.6)	23 (1.6)
Public	8 (0.9)	49 (1.9)	18 (1.3)	2 (0.6)	8 (0.9)	15 (1.3)	23 (1.5)
Catholic and Other Private	3 (1.2)	51 (3.0)	18 (2.3)	4 (1.1)	5 (1.4)	18 (2.3)	23 (2.7)

The standard errors of the estimated percentages appear in parentheses. It can be said with about 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details). Percentages may not total 100 percent due to rounding error.

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment

On the other hand, nearly one-fourth (23 percent) of the fourth graders -- 8 percent by using figures and another 16 percent through well-labelled or explained diagrams -- were able to communicate the relationship between the concept of fractional part and relative size. Because the topics of fractions and proportional reasoning may be covered to varying degrees across the fourth-grade curriculum, the relatively uniform success of fourth graders suggests some experience with sharing such things as cookies or pizzas with friends and families or at least some familiarity with the concepts of "bigger than" and "smaller than."

Across the various demographic subpopulations, performance did not differ significantly for the regions, gender, or type of school. However, a greater percentage of White fourth graders (28 percent) provided satisfactory or better responses than did their Black and Hispanic classmates (9 and 12 percent, respectively). Also, students attending schools in advantaged urban communities outperformed those attending schools in the three other types of communities. More than one-third (35 percent) of the students in advantaged urban schools provided satisfactory or extended responses compared to 14 percent of those attending disadvantaged urban schools.

In general, as shown in TABLE 2.2, the performance of public-school fourth graders across the states tended to parallel that of the nation. For 20 of the participating states, an estimated from one-fifth to one-fourth of the students provided satisfactory or better responses. Additionally, it was estimated that more than one-fourth of the students provided satisfactory or better responses in Connecticut, Iowa, Maine, Minnesota, Missouri, Nebraska, New Hampshire, and North Dakota. However, for 34 of the jurisdictions, it was estimated that at least one-half of the fourth graders did not show any evidence on this question of being able to communicate mathematical concepts.

TABLE 2.2 | **Percentages for Responses to Extended-Response Question, "Pizza Comparision"**

PUBLIC SCHOOLS	Grade 4 - 1992						
	No Response	Incorrect	Minimal	Partial	Satisfactory	Extended	Satisfactory or Better
NATION	8 (0.9)	49 (1.9)	18 (1.3)	2 (0.6)	8 (0.9)	15 (1.3)	23 (1.5)
Northeast	8 (2.3)	41 (5.0)	19 (3.9)	2 (1.4)	8 (2.4)	22 (4.3)	29 (5.1)
Southeast	7 (1.2)	55 (4.7)	17 (2.3)	1 (0.6)	6 (1.4)	14 (2.2)	20 (2.7)
Central	8 (1.9)	49 (3.5)	18 (2.0)	3 (1.5)	8 (2.1)	14 (2.8)	23 (2.8)
West	8 (1.7)	50 (2.8)	17 (2.2)	2 (1.0)	10 (1.2)	13 (2.0)	23 (2.1)
STATES							
Alabama	5 (0.8)	57 (2.9)	18 (2.2)	3 (0.7)	4 (0.8)	12 (1.7)	16 (1.7)
Arizona	5 (1.1)	56 (2.2)	18 (1.5)	2 (0.6)	6 (1.0)	12 (1.3)	19 (1.6)
Arkansas	4 (0.9)	57 (2.3)	17 (1.7)	2 (0.7)	6 (1.2)	14 (1.5)	20 (1.7)
California	10 (1.5)	55 (2.7)	18 (1.5)	2 (0.7)	3 (0.8)	11 (1.7)	14 (1.9)
Colorado	5 (0.8)	51 (2.1)	19 (1.7)	4 (0.6)	6 (0.7)	16 (1.3)	21 (1.5)
Connecticut	7 (0.9)	46 (2.0)	17 (1.6)	3 (0.7)	5 (1.0)	23 (1.5)	27 (1.9)
Delaware	4 (1.0)	58 (2.6)	16 (1.8)	1 (0.6)	7 (1.4)	14 (1.2)	21 (1.5)
Dist. Columbia	10 (1.3)	63 (1.9)	13 (1.5)	1 (0.6)	6 (1.0)	7 (1.3)	12 (1.6)
Florida	6 (1.1)	58 (2.3)	16 (1.6)	3 (0.8)	4 (0.8)	13 (1.3)	17 (1.4)
Georgia	4 (0.9)	52 (2.2)	17 (1.7)	2 (0.6)	8 (1.2)	17 (1.4)	25 (1.7)
Hawaii	6 (1.1)	56 (2.3)	17 (1.9)	3 (0.8)	6 (1.1)	11 (1.3)	17 (1.8)
Idaho	6 (1.1)	50 (2.1)	19 (1.5)	3 (0.8)	7 (0.9)	16 (1.8)	23 (2.0)
Indiana	4 (0.7)	50 (3.0)	19 (2.0)	4 (0.8)	6 (1.1)	18 (1.9)	24 (2.3)
Iowa	3 (0.9)	48 (2.1)	17 (1.6)	3 (0.5)	8 (1.3)	21 (1.8)	29 (1.6)
Kentucky	4 (0.9)	57 (2.6)	15 (1.8)	3 (0.6)	5 (1.0)	16 (1.8)	21 (2.0)
Louisiana	7 (1.1)	61 (2.2)	16 (1.7)	2 (0.7)	5 (0.9)	9 (1.4)	14 (1.5)
Maine	4 (0.9)	40 (2.7)	22 (2.2)	4 (1.1)	9 (1.3)	21 (2.3)	30 (2.3)
Maryland	5 (1.0)	48 (1.8)	23 (2.0)	3 (0.6)	5 (0.9)	16 (1.4)	21 (1.6)
Massachusetts	4 (0.8)	50 (2.9)	21 (2.3)	2 (0.6)	5 (1.1)	17 (2.2)	22 (2.3)
Michigan	5 (1.0)	52 (2.7)	19 (1.9)	3 (0.6)	7 (1.3)	14 (1.7)	21 (1.8)
Minnesota	4 (0.9)	51 (2.4)	16 (1.7)	3 (0.6)	6 (1.1)	21 (2.0)	27 (2.0)
Mississippi	6 (1.1)	65 (2.1)	16 (1.7)	2 (0.6)	3 (0.6)	8 (1.2)	11 (1.3)
Missouri	4 (0.9)	50 (2.1)	17 (1.8)	2 (0.7)	7 (1.1)	19 (1.8)	26 (2.1)
Nebraska	4 (1.0)	49 (2.6)	18 (1.8)	3 (0.7)	7 (1.3)	19 (2.5)	26 (2.5)
New Hampshire	5 (1.2)	42 (2.7)	22 (2.4)	4 (0.8)	5 (1.0)	23 (1.8)	28 (2.2)
New Jersey	5 (0.9)	48 (2.1)	19 (2.0)	5 (0.9)	4 (0.8)	18 (1.9)	22 (1.7)
New Mexico	10 (1.5)	56 (2.6)	15 (1.7)	3 (0.7)	6 (1.0)	10 (1.3)	17 (1.4)
New York	7 (1.2)	56 (2.2)	18 (2.1)	3 (0.7)	5 (0.9)	11 (1.6)	16 (1.7)
North Carolina	4 (0.8)	56 (2.0)	17 (1.7)	3 (0.7)	8 (1.2)	12 (1.2)	19 (1.5)
North Dakota	3 (0.8)	46 (2.3)	18 (2.0)	3 (0.9)	6 (1.1)	24 (2.0)	30 (2.0)
Ohio	2 (0.7)	56 (2.3)	16 (1.6)	3 (0.6)	5 (0.9)	19 (1.5)	24 (1.8)
Oklahoma	4 (0.9)	52 (2.4)	18 (1.8)	3 (0.8)	7 (1.3)	16 (1.7)	23 (1.8)
Pennsylvania	5 (0.6)	53 (2.4)	16 (1.6)	2 (0.6)	6 (1.1)	18 (1.7)	24 (1.7)
Rhode Island	7 (1.3)	52 (2.3)	16 (1.8)	4 (0.8)	6 (1.2)	16 (1.7)	22 (2.1)
South Carolina	3 (0.8)	60 (2.1)	17 (1.6)	3 (0.7)	5 (0.9)	12 (1.3)	17 (1.6)
Tennessee	5 (1.0)	53 (2.4)	16 (1.6)	3 (0.7)	7 (1.3)	16 (1.6)	23 (2.1)
Texas	5 (1.0)	58 (2.4)	16 (1.7)	4 (0.9)	5 (1.0)	12 (1.5)	17 (1.9)
Utah	7 (1.1)	51 (2.0)	17 (1.8)	3 (0.6)	8 (1.1)	15 (1.4)	23 (1.8)
Virginia	6 (1.1)	50 (2.2)	18 (1.5)	3 (0.7)	8 (1.0)	15 (1.7)	24 (1.6)
West Virginia	6 (1.1)	56 (2.4)	16 (1.9)	3 (0.7)	5 (0.9)	14 (1.4)	19 (1.7)
Wisconsin	4 (0.8)	47 (2.1)	21 (1.4)	3 (0.6)	8 (1.1)	17 (1.5)	25 (1.7)
Wyoming	4 (0.7)	52 (2.3)	13 (1.3)	5 (1.0)	6 (1.2)	19 (1.5)	25 (1.9)
TERRITORY							
Guam	10 (1.6)	67 (2.7)	14 (2.2)	2 (0.8)	4 (1.1)	3 (0.7)	7 (1.4)

The standard errors of the estimated percentages appear in parentheses. It can be said with 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details). Percentages may not total 100 percent due to rounding error.

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment.

Grade 4 Question: Laura's Calculator Correction

The Task

Think carefully about the following question. Write a complete answer. You may use drawings, words, and numbers to explain your answer. Be sure to show all of your work.

Laura wanted to enter the number 8375 into her calculator. By mistake, she entered the number 8275. Without clearing the calculator, how could she correct her mistake?

Without clearing the calculator, how could she correct her mistake another way?

Did you use the calculator on this question?

Yes No

Possible Solution

Laura could add 100 to the number in the display because she needs to increase the digit in the hundreds' place by 1.

She also could add 50 two times or 25 four times, or add 1,000 and subtract 900.

She also could describe any other series of arithmetic operations that yields 8375.

Students with a minimal understanding have essentially cleared the calculator by means other than using the **on/c** , **c**, or **ce** buttons. They have demonstrated no understanding of place value. Students at a higher level are beginning to understand place value, but may have focused on the tens' or the thousands' place, rather than the hundreds' place. For a complete response, it is critical that students realize 100 needs to be added to 8275 in order for the calculator screen to display 8375. This can be done without clearing the calculator either directly by the addition of 100 or by performing a series of appropriate arithmetic operations (such as adding ten 10's or by subtracting 1900 and adding 2000) that results in the addition of 100.

National Results, Scoring Guide, and Sample Responses

National Percent for Each Category*	Rating and Performance Category	
17 (1.3)	0	No Response
44 (1.6)	1	Incorrect -- The work is completely incorrect or irrelevant, or the response states, "I don't know."

This INCORRECT response is irrelevant since it ignored the condition given in the problem – that the calculator could not be cleared.

Clear and stat all over.

National Percent for Each Category		Rating and Performance Category
9 (0.8)	2	**Minimal -- Student's response involves attaining a display of 0 on the calculator with a method other than using the on/c, c, or ce keys. Responses in this category demonstrate no connection between place value and arithmetic operations in this problem setting.**

This is a MINIMAL response in which the student was able to obtain a 0 on the calculator display without the use of the on/c, c, or ce keys on the calculator.

Subtract the number she made a mistake on, and she'll get the number 0 and it's clear.

10 (1.1)	3	**Partial -- Student's response begins to connect place value and arithmetic operations as both being necessary to change 8275 to 8375 without clearing the calculator. Errors in arithmetic and/or understanding are evident.**

This is a PARTIAL response, in which the student realizes a 1 must be added to the 2, but makes a place value error.

She could add 1.

She could add 2 and subtract 1.

* The standard errors of the estimated percentages appear in parentheses.

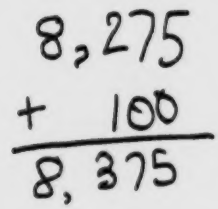
National Percent for Each Category

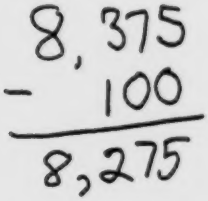
Rating and Performance Category

13 (1.2)

4 **Satisfactory -- Student's response describes only one correct way to change 8275 to 8375.**

This SATISFACTORY response shows one clear method that corrects the place value mistake without clearing the calculator.





7 (0.9)

5 **Extended -- Student's response describes two correct ways to change 8275 to 8375.**

This EXTENDED response shows two different ways to correct the place value mistake without the need to clear the calculator.

She could add 100 more

If she subtracted 100 she could add 200.

Performance Highlights: Laura's Calculator Correction

As suggested by the results presented in TABLE 2.3, the overwhelming majority of students, and perhaps for good reason, found it difficult to accept the premise of this question, pointing out some of the difficulty in developing these types of questions or activities. Because place value and basic number facts receive heavy emphasis in the mathematics curriculum at and prior to grade 4, expectations would be for relatively high performance on this question. However, students had difficulty in accepting another approach to Laura's situation than using the clear button. Nationally, 17 percent of the students left their booklets blank for this question, 44 percent recommended that Laura clear the calculator anyway, and another 9 percent essentially showed her how to clear the calculator without using the clear button. Ten percent appeared to have accepted the premise, but were unable to provide a response that would help rectify Laura's error in entering 8275 instead of 8375.

TABLE 2.3 National Results for Demographic Subgroups for the Extended-Response Question "Laura's Calculator Correction"

Grade 4

	No Response	Incorrect	Minimal	Partial	Satisfactory	Extended	Satisfactory or Better
Nation	17 (1.3)	44 (1.6)	9 (0.8)	10 (1.1)	13 (1.2)	7 (0.9)	20 (1.5)
Northeast	19 (2.7)	34 (3.7)	12 (2.9)	9 (2.5)	14 (2.0)	12 (3.0)	25 (3.9)
Southeast	20 (3.3)	48 (2.3)	9 (1.6)	8 (1.0)	10 (2.2)	5 (1.7)	15 (2.8)
Central	13 (2.2)	47 (4.5)	6 (1.1)	11 (1.8)	17 (3.5)	6 (1.3)	23 (3.4)
West	18 (2.0)	44 (1.9)	9 (1.2)	12 (2.8)	12 (1.7)	6 (1.6)	18 (2.3)
White	15 (1.4)	41 (2.1)	10 (1.1)	10 (1.3)	16 (1.8)	9 (1.3)	24 (2.2)
Black	28 (3.1)	50 (4.2)	6 (1.3)	12 (2.7)	4 (1.7)	1 (0.5)	5 (1.8)
Hispanic	20 (2.7)	56 (2.5)	6 (1.3)	7 (1.8)	8 (1.6)	3 (1.1)	11 (2.2)
Male	19 (2.0)	42 (1.8)	9 (1.0)	8 (1.1)	13 (1.6)	8 (1.3)	22 (1.8)
Female	15 (1.4)	46 (2.2)	9 (1.1)	12 (1.6)	13 (1.5)	5 (1.[illegible])	18 (2.0)
Advantaged Urban	12 (2.1)	34 (4.3)	10 (2.5)	12 (2.7)	21 (3.0)	11 (2.7)	32 (4.8)
Disadvantaged Urban	32 (3.5)	50 (3.3)	6 (2.1)	8 (2.6)	4 (1.5)	0 (0.0)	4 (1.5)
Extreme Rural	22 (4.9)	42 (8.0)	11 (2.2)	6 (1.5)	16 (6.7)	3 (1.9)	20 (6.3)
Other	15 (1.3)	45 (1.5)	9 (0.9)	11 (1.5)	12 (1.3)	8 (1.1)	20 (1.7)
Public	18 (1.4)	45 (1.7)	9 (1.0)	10 (1.3)	12 (1.4)	6 (1.0)	19 (1.6)
Catholic and Other Private	14 (1.5)	38 (2.7)	9 (1.3)	10 (1.6)	18 (2.6)	10 (1.4)	28 (2.6)

The standard errors of the estimated percentages appear in parentheses. It can be said with about 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details). When the proportion of students is either 0 percent or 100 percent, the standard error is inestimable. However, percentages 99.5 percent and greater were rounded to 100 percent and percentages or less were rounded to 0 percent. Percentages may not total 100 percent due to rounding error.

SOURCE: National Assessment of Educational Progress (NAEP), 1992, Mathematics Assessment

Nevertheless, about one-fifth of the fourth graders provided Laura with advice consistent with the demands of the task, providing at least one way to rectify the error without clearing the calculator (usually telling her to add 100). For some demographic subgroups, approximately one-fourth of the students provided answers rated as satisfactory or better, including students in the Northeast (25 percent) and Central (23 percent) regions, White students (24 percent), and those attending private schools (28 percent). Nearly one-third (32 percent) of the students attending schools in advantaged urban areas provided satisfactory or better responses to this question. Although regional differences were not significant statistically, those among racial/ethnic groups, between public- and private-school students, and between advantaged and disadvantaged urban students were significantly different.

As shown in TABLE 2.4, similar to the Pizza Comparison question, there was variation across the states on this task, with from 8 to 31 percent of the students providing satisfactory or better responses. Nine states had an estimated one-fourth or more of their students' responses rated as satisfactory or better, including Connecticut, Idaho, Iowa, Maine, Massachusetts, Minnesota, New Hampshire, New Jersey, and Wisconsin.

TABLE 2.4 | **Percentages for Responses to Extended-Response Question, "Laura's Calculator Correction"**

PUBLIC SCHOOLS	Grade 4 - 1992						
	No Response	Incorrect	Minimal	Partial	Satisfactory	Extended	Satisfactory or Better
NATION	17 (1.4)	45 (1.7)	9 (1.0)	10 (1.3)	13 (1.4)	6 (1.0)	19 (1.6)
Northeast	20 (3.0)	34 (4.0)	12 (3.2)	9 (3.1)	13 (2.2)	11 (3.2)	24 (4.3)
Southeast	20 (3.9)	49 (2.5)	10 (1.9)	8 (1.1)	9 (2.6)	4 (1.8)	13 (3.4)
Central	13 (2.4)	48 (4.6)	5 (1.3)	10 (2.3)	18 (4.1)	5 (1.5)	23 (3.7)
West	18 (2.1)	45 (1.9)	9 (1.4)	12 (3.0)	10 (1.9)	6 (1.8)	16 (2.2)
STATES							
Alabama	13 (1.6)	53 (2.2)	8 (1.2)	10 (1.2)	12 (1.7)	4 (0.9)	16 (1.9)
Arizona	14 (1.7)	56 (2.1)	8 (1.1)	7 (1.0)	10 (1.2)	5 (1.0)	15 (1.2)
Arkansas	13 (1.6)	58 (2.1)	12 (1.3)	7 (1.1)	7 (1.3)	2 (0.7)	10 (1.3)
California	20 (2.1)	49 (2.3)	7 (1.3)	7 (1.0)	12 (1.5)	6 (1.3)	18 (1.8)
Colorado	16 (1.6)	44 (2.1)	8 (1.1)	9 (1.1)	14 (1.5)	9 (1.4)	23 (1.7)
Connecticut	11 (1.5)	41 (2.3)	9 (1.4)	8 (1.1)	20 (1.8)	11 (1.3)	30 (2.4)
Delaware	18 (1.4)	43 (2.4)	10 (1.1)	9 (1.3)	13 (1.5)	8 (1.1)	20 (1.6)
Dist. Columbia	25 (1.8)	55 (2.3)	8 (1.5)	4 (0.9)	6 (0.9)	3 (0.5)	9 (1.0)
Florida	15 (1.4)	53 (2.1)	9 (1.0)	8 (1.3)	11 (1.3)	5 (0.9)	16 (1.6)
Georgia	16 (1.5)	52 (1.9)	9 (1.0)	8 (1.1)	10 (1.2)	5 (0.8)	16 (1.5)
Hawaii	16 (1.9)	44 (2.4)	14 (1.5)	7 (1.0)	12 (1.6)	7 (1.1)	20 (1.9)
Idaho	17 (1.5)	41 (2.0)	8 (1.1)	9 (1.3)	17 (1.7)	8 (1.1)	25 (1.9)
Indiana	13 (1.3)	48 (1.8)	9 (1.2)	10 (1.0)	15 (1.5)	6 (0.9)	21 (1.7)
Iowa	9 (1.3)	44 (2.4)	11 (1.5)	8 (1.2)	18 (1.7)	10 (1.1)	28 (2.2)
Kentucky	10 (1.2)	54 (2.2)	10 (1.4)	8 (1.2)	12 (1.4)	6 (0.8)	18 (1.7)
Louisiana	21 (2.2)	55 (2.7)	6 (1.1)	9 (1.1)	5 (1.1)	4 (1.1)	9 (1.5)
Maine	14 (1.9)	37 (3.0)	7 (1.4)	11 (1.7)	18 (2.4)	13 (1.9)	31 (2.9)
Maryland	17 (1.9)	45 (2.0)	9 (1.3)	7 (0.9)	14 (1.3)	8 (1.2)	23 (1.8)
Massachusetts	15 (1.6)	42 (2.8)	7 (1.1)	8 (1.4)	17 (1.8)	10 (1.5)	27 (2.5)
Michigan	13 (1.6)	50 (2.2)	8 (1.1)	8 (1.2)	13 (1.7)	9 (1.4)	22 (2.3)
Minnesota	12 (1.6)	46 (2.4)	8 (1.4)	8 (1.4)	15 (1.5)	11 (1.4)	26 (2.3)
Mississippi	16 (1.6)	59 (2.0)	9 (1.3)	6 (0.8)	8 (1.3)	2 (0.6)	10 (1.3)
Missouri	15 (1.5)	45 (2.2)	10 (1.2)	7 (1.2)	14 (1.5)	9 (1.5)	23 (2.0)
Nebraska	15 (2.1)	44 (2.0)	11 (1.4)	7 (1.0)	15 (1.8)	8 (1.3)	23 (1.9)
New Hampshire	13 (1.1)	38 (2.2)	9 (1.3)	11 (1.5)	19 (2.0)	10 (1.1)	29 (2.0)
New Jersey	13 (1.5)	41 (2.1)	11 (1.4)	8 (1.2)	16 (1.7)	11 (1.5)	27 (2.0)
New Mexico	16 (2.0)	52 (3.2)	6 (1.2)	8 (1.0)	11 (1.2)	7 (2.2)	18 (2.7)
New York	15 (2.0)	53 (2.3)	5 (1.0)	7 (1.1)	14 (1.4)	6 (1.1)	20 (1.8)
North Carolina	18 (2.0)	48 (2.0)	10 (1.3)	8 (1.0)	11 (1.4)	5 (1.0)	16 (1.7)
North Dakota	9 (1.2)	46 (2.5)	12 (1.9)	9 (1.3)	15 (1.4)	9 (1.1)	24 (1.7)
Ohio	11 (1.3)	50 (2.0)	8 (1.1)	8 (1.1)	14 (1.3)	8 (1.2)	23 (1.5)
Oklahoma	11 (1.5)	52 (2.0)	9 (1.2)	8 (1.3)	15 (1.8)	5 (1.0)	21 (1.8)
Pennsylvania	11 (1.4)	46 (1.9)	12 (1.5)	8 (1.1)	15 (1.3)	8 (1.1)	23 (1.6)
Rhode Island	16 (1.9)	43 (2.3)	13 (1.8)	9 (1.3)	14 (1.5)	5 (0.9)	18 (1.7)
South Carolina	13 (1.4)	59 (1.8)	6 (0.9)	9 (1.1)	9 (1.2)	5 (1.1)	13 (1.4)
Tennessee	12 (1.5)	54 (2.2)	9 (1.3)	9 (1.4)	12 (1.8)	4 (0.9)	16 (2.2)
Texas	13 (1.5)	47 (2.5)	11 (1.3)	8 (1.1)	14 (1.4)	8 (1.2)	21 (2.0)
Utah	14 (1.6)	49 (2.3)	8 (1.1)	8 (1.2)	14 (1.6)	6 (1.1)	21 (1.7)
Virginia	16 (1.5)	45 (2.3)	9 (1.4)	7 (1.2)	13 (1.3)	10 (1.4)	23 (1.9)
West Virginia	14 (1.6)	51 (2.6)	9 (1.2)	9 (1.3)	12 (1.4)	5 (1.1)	17 (1.7)
Wisconsin	12 (1.2)	39 (2.0)	11 (1.1)	10 (1.2)	16 (1.8)	13 (2.0)	29 (2.2)
Wyoming	12 (1.4)	46 (2.1)	9 (1.2)	9 (1.3)	15 (1.6)	9 (1.4)	24 (1.9)
TERRITORY							
Guam	19 (2.3)	60 (2.8)	8 (1.4)	5 (0.9)	6 (1.3)	2 (0.6)	8 (1.5)

The standard errors of the estimated percentages appear in parentheses. It can be said with 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details). Percentages may not total 100 percent due to rounding error.

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment.

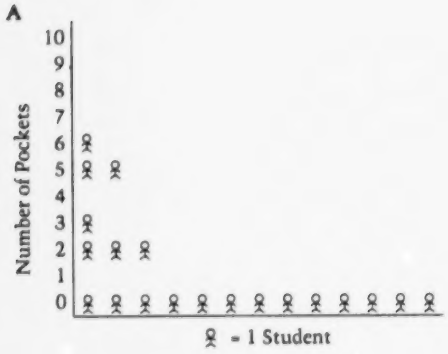
Grade 4 Question: Graphs of Pockets

The Task

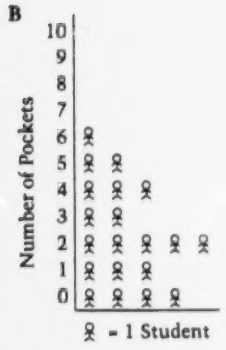
Think carefully about the following question. Write a complete answer. You may use drawings, words, and numbers to explain your answer. Be sure to show all of your work.

There are 20 students in Mr. Pang's class. On Tuesday most of the students in the class said they had pockets in the clothes they were wearing.

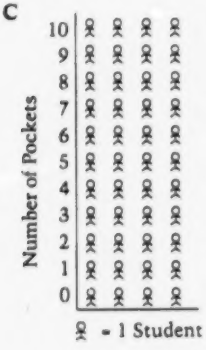
A



B



C



The Task (continued)

Which of the graphs most likely shows the number of pockets that each child had? _____

Explain why you chose that graph.

Explain why you did not choose the other graphs.

Possible Solution

Graph B, because it had 20 students and most of the students have clothes with pockets, or the distribution of the number of pockets is reasonable.

It could not be Graph A because most of the students do not have pockets.

It could not be Graph C since there are more than 20 students shown.

OR, it is not likely that there would be the same number of students for each number of pockets.

OR, most clothes don't have 10 pockets.

Students need to understand the information provided in the question in order to study and determine which graphical representation most accurately depicts the given data and why the other graphs are inappropriate. The essential facts that students need to comprehend are:

- There are 20 students in Mr. Pang's class. (Thus, Graph C is inappropriate because more than 20 students are represented. Additionally, the distribution of the number of pockets is unreasonable.)

- Most students in Mr. Pang's class have clothes with pockets. (Thus, Graph A is inappropriate because most of the 20 students have clothes that do not have pockets.)

Therefore, in reviewing the graphs, only Graph B reasonably conveys the given information accurately since 20 students are represented and most of these students have clothes with pockets. In extended solutions to this task, students must clearly communicate a rationale for the graph they select and explain why the other graphs are inappropriate.

National Results, Scoring Guide, and Sample Responses

National Percent for Each Category*	Rating and Performance Category	
6 (0.7)	0	No Response
46 (1.4)	1	Incorrect -- The work is completely incorrect or irrelevant, or the response states, "I don't know."

This is an INCORRECT response. The student may have picked Graph C because of its rectangular shape. This is inappropriate because it does not use any of the information given about the number of students in the class or that most students had clothes with pockets.

C

because they are all equal

because they are not equal

23 (1.2)	2	Minimal -- Student chooses Graph B with no explanation or the student chooses Graph A or Graph C with an explanation that shows some understanding.

This is a MINIMAL solution since the student did not select Graph B, the most appropriate graph to display the data, but did give a reason that showed some understanding.

A

I did because A had 20 Students.

I did because they had more than 20 students.

* The standard errors of the estimated percentages appear in parentheses.

National Percent for Each Category	Rating and Performance Category
15 (0.8)	3 **Partial -- Student chooses Graph B but does not give an adequate explanation or student chooses Graph B but gives no explanation why, but explains why the answer is neither Graph A nor Graph C.**

This is a PARTIAL response. The student did select the most appropriate graph, B, but did not give a complete explanation of why neither Graph A nor Graph C was the best choice.

B

I chose B because most of the people had pockets.

I didn't choose the other graphs because more people didn't have pockets.

7 (0.7)	4 **Satisfactory -- Student chooses Graph B and gives a good explanation but does not mention the other graphs, or student gives a good explanation of why the answer cannot be Graph A or Graph C, but does not give a good explanation of why the answer is Graph B.**

This is a SATISFACTORY response since the student did select Graph B and also provided a complete explanation by indicating there were 20 people and most of them had pockets. However, the student did not provide any information about why Graphs A and C were inappropriate.

B

B had a total of 20 and not that many people had a pockets

National Percent for Each Category	Rating and Performance Category	
3 (0.6)	5	**Extended -- Student chooses Graph B, explains why the answer must be Graph B, and explains why neither Graph A nor Graph C can be the correct solution.**

This is an EXTENDED response. The student selects Graph B and gives a clear and accurate explanation. In like fashion, the student conveys correct and concise reasons for not choosing either Graph A or Graph C.

GB

I chose graph B because I could read it better, and at the top it said that most of the kids had pockets in their clothes. graph A had a hole bunch of kids who didn't have pockets - I think graph B explained it better

I did not chose the other graphs because graph C had too many kids in the graph and graph A had to many kids didn't have pockets in their clothes

BEST COPY AVAILABLE

Performance Highlights: Graphs of Pockets

The national results for the Graphs of Pockets question presented in TABLE 2.5 indicate that fourth graders had difficulty responding to the many requirements inherent in this task, including assimilating the criteria set forth in the question, using those criteria to evaluate the data in the pictographs, and explaining the results of their analysis. Approximately 46 percent of the fourth graders nationally seemed unfamiliar with this type of assessment task and were unable to relate the pictographs to the problem, as in the example response where the student selected graph C because the number of pockets were equal. Many of the remaining fourth graders tended to perform at the minimal or partial levels. Approximately one-fourth of the students showed a minimal grasp of the relationship between the problem situation and the three graphs, by either selecting the right graph with no explanation or selecting the wrong graph for a reason that was related to Mr. Pang's class. Fifteen percent of the fourth graders provided a partial response based on selecting the right graph and giving some reason to support it via the process of eliminating either graph A or C.

TABLE 2.5 National Results for Demographic Subgroups for the Extended-Response Question, "Graphs of Pockets"

Grade 4

	No Response	Incorrect	Minimal	Partial	Satisfactory	Extended	Satisfactory or Better
Nation	6 (0.7)	46 (1.4)	23 (1.2)	15 (0.8)	7 (0.7)	3 (0.6)	10 (0.9)
Northeast	6 (1.5)	42 (3.7)	27 (2.7)	13 (2.2)	10 (2.1)	2 (0.8)	12 (2.5)
Southeast	6 (1.4)	52 (2.9)	19 (1.8)	14 (1.3)	5 (1.4)	4 (1.4)	9 (2.3)
Central	6 (1.6)	46 (3.1)	22 (3.1)	18 (1.7)	5 (1.2)	4 (1.2)	9 (1.7)
West	6 (1.2)	45 (2.2)	24 (1.8)	15 (1.4)	8 (1.3)	2 (1.0)	10 (1.2)
White	5 (0.8)	42 (1.7)	24 (1.5)	17 (1.1)	9 (1.0)	4 (0.8)	13 (1.3)
Black	6 (1.5)	67 (2.5)	19 (2.1)	7 (1.7)	0 (0.3)	1 (0.2)	1 (0.4)
Hispanic	13 (2.8)	50 (3.4)	23 (3.6)	11 (1.9)	2 (0.9)	0 (0.3)	2 (1.0)
Male	6 (1.0)	47 (2.1)	23 (1.6)	12 (1.2)	7 (1.2)	4 (0.9)	11 (1.5)
Female	5 (0.8)	45 (2.1)	22 (1.8)	18 (1.4)	7 (1.2)	2 (0.6)	9 (1.2)
Advantaged Urban	3 (0.9)	41 (3.5)	22 (2.0)	18 (3.0)	11 (2.7)	5 (1.7)	16 (3.4)
Disadvantaged Urban	9 (2.0)	58 (4.1)	20 (3.0)	10 (2.1)	2 (0.8)	1 (0.5)	2 (0.9)
Extreme Rural	7 (2.6)	46 (5.9)	23 (3.3)	14 (2.4)	8 (1.8)	2 (1.0)	10 (2.3)
Other	6 (0.8)	46 (2.0)	23 (1.6)	15 (1.1)	7 (1.0)	3 (0.7)	10 (1.2)
Public	6 (0.8)	46 (1.6)	23 (1.3)	15 (0.9)	7 (0.8)	3 (0.6)	10 (1.0)
Catholic and Other Private	4 (1.1)	48 (2.4)	20 (1.9)	16 (2.0)	7 (1.5)	4 (1.6)	12 (2.0)

The standard errors of the estimated percentages appear in parentheses. It can be said with about 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details). When the proportion of students is either 0 percent of 100 percent, the standard error is inestimable. However, percentages 99.5 percent and greater were rounded to 100 percent and percentages 0.5 percent or less were rounded to 0 percent. Percentages may not total 100 percent due to rounding error.

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment

Ten percent provided satisfactory or better responses, with 7 percent giving a good explanation for either why graph B could fit the situation described for the number of pockets had by students in Mr. Pang's class, or why both A and C did not. Only 3 percent provided full responses, explaining the pro side for graph B as well as the con sides for both graphs A and C.

Because of the low levels of satisfactory or better performance on this task, in general, percentages of successful performance did not differ significantly across demographic subgroups.

The state-level results for public-school students mirrored those for the nation. The response percentages shown in TABLE 2.6 indicate that for many participating jurisdictions about half the fourth graders or more did not relate the text and graphs for the problem-situation. Of those that showed some understanding, most had difficulty in providing explanations or reasons supporting the fit between the criteria for the students' pockets and the data as displayed in the pictographs. Across the jurisdictions, the percentages of satisfactory or better responses to this task ranged from 3 to 13 percent. Twelve states had an estimated 10 percent or more of their students provide responses judged as satisfactory or extended, and in Connecticut, North Dakota, and Pennsylvania at least 5 percent of the students were estimated to have performed at the extended level.

TABLE 2.6 | **Percentages for Responses to Extended-Response Question, "Graphs of Pockets"**

PUBLIC SCHOOLS	Grade 4 - 1992						
	No Response	Incorrect	Minimal	Partial	Satisfactory	Extended	Satisfactory or Better
NATION	6 (0.8)	46 (1.6)	23 (1.3)	15 (0.9)	7 (0.8)	3 (0.6)	10 (1.0)
Northeast	7 (1.6)	39 (4.0)	29 (2.9)	12 (2.3)	11 (2.1)	2 (0.9)	12 (2.5)
Southeast	6 (1.8)	53 (3.6)	18 (2.1)	14 (1.5)	4 (1.4)	4 (1.4)	8 (2.4)
Central	6 (1.8)	45 (3.7)	22 (3.4)	18 (2.0)	5 (1.4)	4 (1.5)	8 (2.1)
West	6 (1.3)	46 (2.6)	24 (1.9)	14 (1.5)	8 (1.4)	2 (1.1)	10 (1.4)
STATES							
Alabama	3 (0.6)	56 (2.1)	21 (2.0)	15 (1.6)	4 (0.7)	1 (0.4)	5 (0.9)
Arizona	5 (1.0)	52 (1.8)	20 (1.5)	16 (1.4)	5 (0.9)	2 (0.6)	7 (1.0)
Arkansas	2 (0.7)	57 (2.2)	19 (1.9)	15 (1.4)	5 (1.0)	2 (0.7)	7 (1.1)
California	5 (1.1)	53 (2.5)	20 (1.5)	15 (1.9)	4 (1.1)	2 (0.6)	6 (1.2)
Colorado	3 (0.7)	44 (2.0)	27 (1.5)	18 (1.4)	6 (1.1)	2 (0.5)	8 (1.2)
Connecticut	4 (0.9)	43 (2.6)	24 (1.8)	17 (1.4)	8 (1.1)	5 (1.0)	12 (1.6)
Delaware	4 (0.9)	53 (2.5)	21 (2.3)	16 (1.7)	5 (1.1)	1 (0.7)	7 (1.1)
Dist. Columbia	7 (1.1)	62 (2.0)	21 (1.8)	6 (1.2)	2 (0.7)	1 (0.4)	3 (0.8)
Florida	5 (1.0)	52 (2.0)	21 (1.8)	16 (1.4)	4 (1.0)	2 (0.5)	6 (1.2)
Georgia	2 (0.5)	50 (2.6)	22 (2.0)	19 (1.8)	4 (0.9)	3 (0.9)	8 (1.2)
Hawaii	5 (0.8)	56 (2.3)	17 (1.7)	15 (1.7)	5 (1.0)	2 (0.6)	7 (1.1)
Idaho	4 (0.6)	44 (2.2)	24 (1.6)	21 (1.9)	5 (0.9)	2 (0.7)	7 (1.0)
Indiana	1 (0.4)	53 (2.3)	23 (1.8)	17 (1.8)	4 (1.1)	2 (0.6)	6 (1.1)
Iowa	3 (0.7)	42 (2.0)	25 (1.7)	17 (1.6)	8 (1.1)	4 (0.8)	12 (1.3)
Kentucky	3 (0.7)	56 (2.4)	20 (1.8)	15 (1.7)	5 (1.1)	1 (0.5)	6 (1.2)
Louisiana	5 (0.9)	57 (2.5)	18 (1.8)	15 (1.7)	3 (0.7)	2 (0.5)	5 (0.8)
Maine	3 (0.7)	38 (1.9)	26 (2.2)	21 (2.1)	8 (1.3)	4 (0.9)	13 (1.6)
Maryland	5 (1.0)	50 (2.2)	19 (1.5)	17 (1.6)	6 (1.0)	3 (0.6)	9 (1.2)
Massachusetts	5 (0.8)	43 (2.2)	24 (1.9)	18 (1.5)	7 (1.2)	3 (0.9)	10 (1.6)
Michigan	1 (0.4)	51 (2.0)	20 (1.6)	18 (1.5)	7 (1.1)	3 (0.6)	10 (1.2)
Minnesota	3 (0.7)	42 (1.8)	22 (1.5)	22 (1.8)	7 (1.2)	4 (0.8)	11 (1.4)
Mississippi	4 (0.8)	60 (2.3)	20 (1.8)	12 (1.3)	3 (0.7)	1 (0.4)	4 (0.8)
Missouri	1 (0.5)	47 (2.2)	25 (1.6)	18 (1.7)	6 (1.1)	3 (0.7)	9 (1.3)
Nebraska	5 (1.0)	42 (2.3)	24 (1.7)	20 (1.9)	6 (1.2)	3 (0.9)	9 (1.1)
New Hampshire	3 (0.8)	46 (2.6)	23 (1.9)	19 (2.0)	6 (1.4)	3 (0.7)	9 (1.5)
New Jersey	3 (0.8)	47 (2.4)	23 (1.9)	18 (2.0)	7 (1.2)	3 (0.8)	10 (1.5)
New Mexico	2 (0.6)	59 (2.2)	21 (1.7)	14 (1.9)	4 (0.9)	1 (0.3)	4 (1.0)
New York	4 (1.3)	49 (2.4)	22 (1.8)	16 (1.6)	5 (1.2)	4 (1.0)	9 (1.5)
North Carolina	3 (0.6)	52 (2.2)	24 (1.5)	15 (1.5)	3 (0.6)	3 (0.6)	6 (0.9)
North Dakota	4 (1.1)	44 (2.6)	22 (1.8)	20 (1.9)	6 (1.0)	5 (1.0)	10 (1.1)
Ohio	4 (0.9)	51 (2.4)	20 (1.5)	15 (1.6)	6 (1.0)	4 (0.9)	9 (1.4)
Oklahoma	2 (0.7)	46 (2.2)	27 (1.8)	18 (1.8)	3 (0.9)	3 (0.7)	7 (1.0)
Pennsylvania	4 (0.8)	49 (2.1)	22 (2.0)	15 (1.6)	6 (1.2)	5 (0.9)	11 (1.7)
Rhode Island	6 (1.1)	52 (2.3)	22 (1.9)	14 (1.6)	5 (1.1)	2 (0.7)	7 (1.4)
South Carolina	2 (0.5)	52 (2.4)	23 (1.9)	18 (1.7)	3 (0.6)	2 (0.4)	6 (0.8)
Tennessee	3 (0.7)	55 (2.2)	20 (1.7)	16 (1.5)	4 (0.9)	2 (0.6)	6 (1.1)
Texas	3 (0.8)	47 (2.6)	19 (1.8)	20 (1.5)	8 (1.2)	4 (1.2)	11 (1.5)
Utah	3 (0.8)	47 (2.1)	27 (1.6)	17 (1.2)	4 (0.7)	3 (0.6)	6 (1.0)
Virginia	3 (0.4)	50 (2.1)	21 (1.4)	18 (1.6)	6 (1.1)	2 (0.6)	8 (1.6)
West Virginia	3 (0.6)	57 (2.0)	20 (1.7)	15 (1.6)	3 (0.8)	2 (0.7)	5 (1.0)
Wisconsin	3 (0.7)	48 (2.5)	20 (1.9)	18 (1.6)	8 (1.1)	3 (0.8)	11 (1.3)
Wyoming	4 (0.8)	47 (2.2)	23 (2.2)	15 (1.6)	8 (1.1)	3 (0.7)	11 (1.2)
TERRITORY							
Guam	6 (1.1)	66 (2.0)	15 (1.5)	10 (1.5)	2 (0.5)	2 (0.6)	3 (0.7)

The standard errors of the estimated percentages appear in parentheses. It can be said with 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details). Percentages may not total 100 percent due to rounding error.

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment.

Extended-Response Questions: Grade 8

The eighth-grade students participating in NAEP's 1992 mathematics assessment were given six extended-response tasks, three of which are presented in the following section.

"Treena's Budget," asking students to select options that fit an overall budget, was designed to be accessible to all eighth-grade students. They were required to use basic whole number operations skills to figure out which set of travel and instructional alternatives was available to Treena for basketball camp. They were given a scientific calculator (TI-30) to use if they wished.

In "Radio Stations," students were asked to apply measurement and geometry knowledge to diagram the intersection of signal transmissions from two radio stations.

"Marcy's Dot Pattern" is a prealgebra question requiring pattern recognition and an elementary understanding of the concept of recursion. Students were asked to generalize about any term in a pattern of dots and use their generalization to extend the pattern. Students also were provided the scientific calculator to use in solving this problem.

Grade 8 Question: Treena's Budget

The Task

This question requires you to show your work and explain your reasoning. You may use drawings, words, and numbers in your explanation. Your answer should be clear enough so that another person could read it and understand your thinking. It is important that you show all your work.

Treena won a 7-day scholarship worth $1,000 to the Pro Shot Basketball Camp. Round-trip travel expenses to the camp are $335 by air or $125 by train. At the camp she must choose between a week of individual instruction at $60 per day or a week of group instruction at $40 per day. Treena's food and other expenses are fixed at $45 per day. If she does not plan to spend any money other than the scholarship, what are all choices of travel and instruction plans that she could afford to make? Explain your reasoning.

Did you use the calculator on this question?

Yes No

Possible Solution

The solution to this task requires students to use everyday consumer sense to determine Treena's fixed expenses and analyze the various choices she has for travel (plane or train) and instruction (individual or group). Students also must compare the total cost for **each** of the four alternatives to which this analysis leads to the $1,000 value of Treena's scholarship, in order to conclude which choices meet the given conditions.

Treena's fixed expenses will be $45 x 7 = $315 for the seven days. Therefore, she has $1,000 – $315 = $685 to spend on travel and instruction. Travel costs are either train ($125) or plane ($335). Instruction costs are either group ($40 x 7 = $280), or individual ($60 x 7 = $420).

The four choices Treena has are:

Travel by train, group instruction, and fixed expenses: $125 + $280 + $315 = $720

Travel by plane, group instruction, and fixed expenses: $335 + $280 + $315 = $930

Possible Solution (continued)

Travel by train, individual instruction, and fixed expenses: $125 + $420 + $315 = $860

Travel by plane, individual instruction, and fixed expenses: $335 + $420 + $315 = $1,070

Students must realize that Treena cannot choose the individual plan and travel by plane because the total expenses ($1,070) would be greater than the allotted scholarship. Any full credit response must clearly communicate that Treena has three options that do not exceed $1,000, what the three options are, and how the student arrived at the three options.

National Results, Scoring Guide, and Sample Responses

National Percent for Each Category*	Rating and Performance Category	
22 (1.2)	0	**No Response**
37 (1.6)	1	**Incorrect -- The work is completely incorrect or irrelevant, or the response states, "I don't know."**

This INCORRECT response appears to be somewhat on task but the work shown does not warrant credit even at the minimal level.

Add everything other then scholarship and you will get 230.

* The standard errors of the estimated percentages appear in parentheses.

National Percent for Each Category	Rating and Performance Category	
22 (1.2)	2	**Minimal -- a) Student indicates one or more options only (such as group and train) with no supporting evidence, or b) Student work contains major mathematical errors and/or flaws in reasoning (for example, the student does not consider Treena's fixed expenses).**

This MINIMAL response does illustrate one valid budget option, but does not show any supporting calculations.

She could take the train to camp have individual instruction and eat everyday and not run out of money.

15 (1.0)	3	**Partial -- The student a) indicates one or more correct options; additional supporting work beyond the minimal level must be present, but the work may contain some computational errors; or b) demonstrates correct mathematics for one or two options, but does not indicate the options that are supported by his or her mathematics.**

This PARTIAL response illustrates one acceptable budget alternative (group and train) and the corroborating computational work.

train at $ 45

40 × 7 = 280

280 + 315 = 720

group at

345 × 7 = 315

$720 would she all spend

She just took the cheapest ones of her choices now she has money left over

National Percent for Each Category	Rating and Performance Category	
2 (0.3)	4	**Satisfactory -- The student a) shows correct mathematical evidence that Treena has three options, but the supporting work is incomplete; or b) shows correct mathematical evidence for any two of Treena's three options and the supporting work is clear and complete.**

This SATISFACTORY response illustrates two appropriate budget options (both individual and train and group and plane) as well as the correct supporting calculations.

125 + 420 + 315 = $860

$1000 > $860

If $1000 is more than $800 she has money left over so she could take private lessons, a train and her food.

335 + 315 + 280 = 930

$1000 > $930

She could take a plane, her food, and group lessons

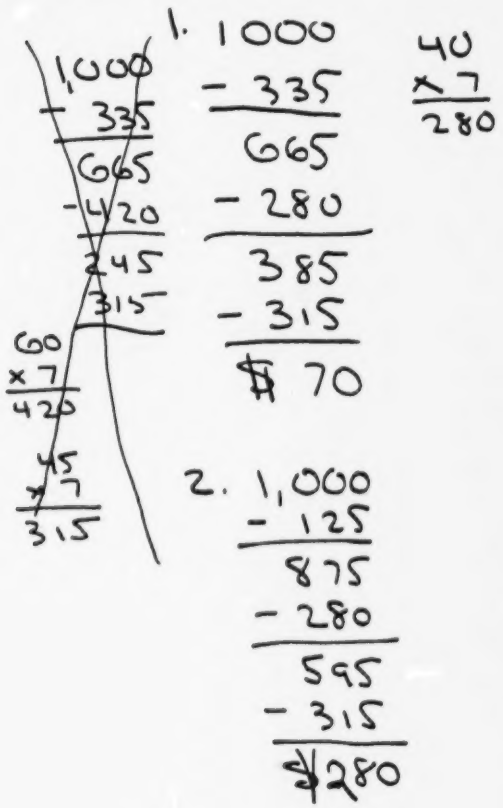
National Percent for Each Category

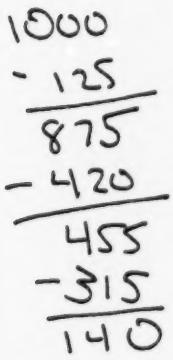
2 (0.4)

This outstanding EXTENDED response provides the correct calculations in terms of the excess dollars that remain from the $1,000 scholarship, for the three acceptable budget options.

Rating and Performance Category

5 **Extended -- The correct solution indicates what the three possible options are and includes supporting work for each option.**





1. take air, group, food
2. train, group, food
3. train, individual, food

Performance Highlights: Treena's Budget

As shown in TABLE 2.7, eighth graders had considerable difficulty understanding and persevering with this problem situation. Because the mathematics involved whole number computations and the context was budgeting, both thought to be familiar to eighth graders, high performance was anticipated. Further, students were given a calculator. However, nationally, more than one-fifth of the students left their papers blank. Another 37 percent could not seem to make the translation from the problem to the required calculations. Many of these responses, as depicted in the illustration, ignored the question asked, instead providing a calculation or two based on numbers in the problem or providing otherwise unrelated information. These students did not communicate that they even understood the set of boundary conditions and how they related to the stated budget constraints.

TABLE 2.7. National Results for Demographic Subgroups for the Extended-Response Task, "Treena's Budget"

Grade 8

	No Response	Incorrect	Minimal	Partial	Satisfactory	Extended	Satisfactory or Better
Nation	22 (1.2)	37 (1.6)	22 (1.2)	15 (1.0)	2 (0.3)	2 (0.4)	4 (0.5)
Northeast	23 (3.0)	37 (3.7)	21 (2.7)	14 (2.0)	2 (0.4)	2 (1.0)	4 (1.0)
Southeast	23 (1.9)	43 (3.0)	22 (2.3)	10 (1.7)	2 (0.6)	1 (0.5)	3 (0.8)
Central	17 (2.0)	34 (3.9)	25 (2.6)	18 (1.6)	4 (0.8)	2 (1.2)	6 (1.1)
West	26 (2.9)	34 (2.3)	21 (2.2)	17 (2.5)	2 (0.7)	1 (0.7)	3 (0.9)
White	18 (1.5)	34 (2.1)	25 (1.6)	18 (1.2)	3 (0.4)	2 (0.6)	5 (0.6)
Black	30 (2.9)	50 (3.4)	13 (2.1)	7 (1.7)	0 (0.5)	0 (0.0)	0 (0.5)
Hispanic	36 (3.3)	40 (3.0)	19 (2.7)	5 (1.3)	1 (0.4)	0 (0.4)	1 (0.6)
Male	28 (1.8)	38 (2.1)	19 (1.5)	12 (1.1)	1 (0.4)	1 (0.4)	2 (0.5)
Female	16 (1.8)	36 (1.9)	26 (1.7)	17 (1.8)	3 (0.6)	3 (0.8)	6 (0.8)
Advantaged Urban	10 (3.2)	33 (3.4)	35 (3.4)	15 (3.0)	3 (1.3)	4 (1.6)	7 (2.7)
Disadvantaged Urban	42 (3.3)	39 (4.1)	10 (2.5)	8 (2.1)	2 (1.1)	0 (0.0)	2 (1.1)
Extreme Rural	24 (5.3)	30 (5.1)	23 (3.9)	19 (4.1)	2 (1.3)	3 (1.7)	4 (1.8)
Other	21 (1.4)	38 (2.0)	22 (1.5)	15 (1.4)	2 (0.5)	2 (0.5)	4 (0.6)
Public	23 (1.4)	37 (1.8)	21 (1.3)	14 (1.1)	2 (0.4)	2 (0.5)	4 (0.5)
Catholic and Other Private	14 (2.3)	35 (2.7)	30 (2.4)	16 (2.2)	2 (0.8)	2 (0.7)	4 (1.0)

The standard errors of the estimated percentages appear in parentheses. It can be said with about 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details). When the proportion of students is either 0 percent or 100 percent, the standard error is inestimable. However, percentages 99.5 percent and greater were rounded to 100 percent and percentages 0.5 percent or less were rounded to 0 percent. Percentages may not total 100 percent due to rounding error.

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment

More than one-third of the students received either minimal (22 percent) or partial (15 percent) credit for their work. These students appeared to have understood the mathematics and the question, but failed to heed the instructions asking that work be shown, reasoning explained, and multiple budget options explored. Only 4 percent of the students provided fully documented work -- explanations and supporting computations -- for at least two of the three viable budget options.

Across the subgroups of students, the highest percentage of satisfactory or better responses appeared to have been posted by advantaged urban students -- 7 percent, although this result was not statistically different from the percentages attained by students in any other types of communities. The very low performance overall precluded substantial differences among subgroups in levels of successful performance. However, 30 percent or more of the Black and Hispanic students left this question blank in comparison to 18 percent of their White counterparts. Similarly, 42 percent of the disadvantaged urban students left this question blank compared to 10 percent of the advantaged urban students. It may be that in learning to engage in more complex assessment scenarios, students need to be encouraged to spend more time simply digesting the question. Also, it may be that the communication aspects of reading and listening need to be emphasized as well as their counterparts of writing and discussing. Because from an estimated one-third to half the students across the demographic groups provided incorrect and unrelated information, the results suggest that some students may rush to provide an answer through meaningless manipulations without thinking through problems and strategies for solving them.

Given that for this question, ratings of minimal or better indicated basic understanding of the mathematics and task (albeit incomplete implementation), it is interesting to note that only nine states had an estimated majority of their students reach this level or better: Connecticut, Iowa, Maine, Minnesota, Nebraska, New Hampshire, North Dakota, Wisconsin, and Wyoming (see TABLE 2.8). The percentages of students providing a satisfactory or better response ranged from 0 to 8 percent. States with 7 to 8 percent of their students estimated to have provided such responses included Colorado, Connecticut, and Iowa.

TABLE 2.8 | **Percentages for Responses to Extended-Response Question, "Treena's Budget"**

PUBLIC SCHOOLS	Grade 8 - 1992						
	No Response	Incorrect	Minimal	Partial	Satisfactory	Extended	Satisfactory or Better
NATION	23 (1.4)	37 (1.8)	21 (1.3)	14 (1.1)	2 (0.4)	2 (0.5)	4 (0.5)
Northeast	23 (3.8)	40 (4.5)	19 (2.7)	13 (2.7)	2 (0.3)	3 (1.3)	5 (1.3)
Southeast	24 (2.2)	43 (3.7)	20 (2.5)	10 (1.8)	2 (0.7)	1 (0.5)	3 (0.9)
Central	18 (2.2)	34 (4.2)	25 (2.9)	18 (1.6)	4 (0.9)	2 (1.3)	6 (1.2)
West	27 (3.1)	34 (2.4)	20 (2.4)	17 (2.5)	2 (0.7)	1 (0.7)	3 (1.0)
STATES							
Alabama	26 (2.2)	40 (2.1)	20 (1.6)	12 (1.6)	2 (0.6)	0 (0.3)	2 (0.7)
Arizona	23 (2.0)	35 (2.0)	23 (1.9)	16 (1.8)	3 (0.8)	1 (0.3)	4 (0.9)
Arkansas	24 (2.0)	38 (1.8)	25 (1.9)	11 (1.2)	2 (0.6)	0 (0.2)	2 (0.6)
California	27 (2.2)	36 (2.5)	22 (1.7)	12 (1.7)	2 (0.8)	0 (0.3)	3 (0.8)
Colorado	18 (1.4)	33 (2.0)	24 (1.8)	19 (1.7)	4 (0.6)	3 (0.8)	7 (0.9)
Connecticut	21 (1.9)	29 (2.3)	23 (1.8)	20 (1.9)	4 (0.9)	3 (0.7)	7 (1.0)
Delaware	24 (2.5)	31 (2.4)	23 (2.0)	17 (1.9)	4 (1.2)	1 (0.4)	4 (1.4)
Dist. Columbia	41 (2.4)	38 (2.4)	14 (1.9)	5 (1.2)	0 (0.4)	1 (0.6)	1 (0.7)
Florida	24 (2.2)	37 (2.3)	24 (2.4)	12 (1.5)	2 (0.6)	2 (0.7)	4 (0.9)
Georgia	24 (1.8)	36 (2.0)	25 (1.9)	13 (1.5)	2 (0.5)	0 (0.2)	2 (0.6)
Hawaii	34 (2.0)	36 (2.0)	16 (1.8)	13 (1.4)	1 (0.4)	0 (0.3)	1 (0.5)
Idaho	19 (1.5)	37 (1.9)	23 (1.7)	17 (2.0)	3 (0.7)	1 (0.4)	4 (0.7)
Indiana	17 (1.8)	35 (2.0)	27 (2.2)	18 (2.1)	2 (0.7)	1 (0.4)	3 (0.8)
Iowa	11 (1.1)	29 (2.2)	30 (2.1)	22 (1.9)	4 (0.7)	4 (0.8)	8 (1.1)
Kentucky	19 (1.6)	32 (1.8)	27 (1.8)	18 (1.4)	2 (0.7)	1 (0.4)	4 (0.9)
Louisiana	35 (2.3)	36 (2.0)	19 (1.9)	8 (1.2)	1 (0.4)	0 (0.2)	1 (0.5)
Maine	17 (1.8)	29 (2.0)	29 (2.3)	20 (1.6)	3 (0.9)	2 (0.6)	5 (1.1)
Maryland	26 (2.3)	30 (2.3)	20 (1.8)	20 (1.9)	3 (0.8)	2 (0.5)	4 (0.9)
Massachusetts	21 (1.9)	31 (2.0)	26 (2.1)	17 (1.8)	3 (0.7)	2 (0.6)	5 (1.0)
Michigan	21 (1.7)	34 (2.3)	25 (1.8)	16 (1.8)	2 (0.7)	1 (0.5)	4 (0.9)
Minnesota	13 (1.5)	34 (2.1)	27 (2.7)	19 (1.7)	4 (0.9)	3 (0.9)	6 (1.4)
Mississippi	29 (2.2)	39 (1.9)	23 (1.9)	8 (1.1)	1 (0.4)	1 (0.3)	1 (0.5)
Missouri	20 (1.8)	37 (2.0)	23 (1.6)	16 (1.7)	2 (0.6)	2 (0.3)	4 (0.7)
Nebraska	17 (1.7)	30 (2.1)	29 (1.9)	20 (1.6)	4 (0.8)	1 (0.4)	5 (0.9)
New Hampshire	18 (1.3)	31 (2.1)	26 (1.7)	20 (2.1)	3 (0.7)	1 (0.5)	5 (0.8)
New Jersey	21 (2.2)	33 (3.1)	24 (2.0)	17 (1.6)	3 (0.9)	2 (0.7)	5 (1.3)
New Mexico	22 (1.7)	37 (2.0)	26 (2.0)	13 (1.4)	1 (0.5)	0 (0.1)	1 (0.5)
New York	23 (1.9)	36 (2.4)	21 (1.9)	15 (1.6)	2 (0.7)	2 (0.6)	4 (1.0)
North Carolina	24 (2.1)	35 (2.0)	25 (1.7)	14 (1.4)	2 (0.6)	1 (0.3)	3 (0.8)
North Dakota	14 (1.4)	32 (2.4)	29 (2.1)	19 (1.6)	4 (0.9)	2 (0.6)	6 (1.2)
Ohio	21 (1.9)	33 (1.6)	22 (2.0)	19 (1.9)	3 (0.9)	2 (0.6)	5 (1.0)
Oklahoma	16 (1.6)	37 (2.3)	28 (2.0)	16 (1.9)	2 (0.8)	1 (0.4)	3 (0.8)
Pennsylvania	19 (1.5)	35 (2.3)	23 (1.9)	20 (1.6)	2 (0.6)	1 (0.5)	3 (0.8)
Rhode Island	21 (3.2)	33 (2.5)	24 (2.2)	16 (2.2)	3 (0.9)	2 (0.8)	6 (1.1)
South Carolina	26 (1.9)	37 (2.4)	23 (2.0)	11 (1.4)	1 (0.4)	1 (0.5)	2 (0.5)
Tennessee	26 (2.1)	36 (2.1)	21 (1.9)	14 (1.6)	2 (0.6)	1 (0.4)	3 (0.7)
Texas	24 (1.9)	31 (2.4)	28 (2.0)	13 (1.6)	2 (0.6)	2 (0.8)	4 (1.0)
Utah	21 (1.5)	36 (2.0)	23 (1.6)	16 (1.3)	3 (0.6)	1 (0.3)	4 (0.8)
Virginia	18 (1.4)	33 (2.1)	25 (2.0)	19 (1.9)	2 (0.7)	2 (0.6)	4 (0.9)
West Virginia	22 (1.7)	38 (2.1)	25 (1.6)	13 (1.3)	2 (0.6)	0 (0.3)	2 (0.7)
Wisconsin	16 (2.3)	32 (2.3)	27 (2.6)	19 (2.4)	3 (0.9)	3 (0.7)	6 (1.0)
Wyoming	15 (1.5)	29 (2.0)	31 (2.1)	19 (1.9)	3 (0.7)	2 (0.7)	5 (0.9)
TERRITORIES							
Guam	49 (3.1)	33 (2.8)	12 (1.9)	5 (1.5)	0 (0.3)	0 (0.3)	1 (0.4)
Virgin Islands	61 (2.3)	27 (2.9)	10 (2.1)	1 (0.6)	0 (0.0)	0 (0.0)	0 (0.0)

The standard errors of the estimated percentages appear in parentheses. It can be said with 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details). Percentages may not total 100 percent due to rounding error.

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment.

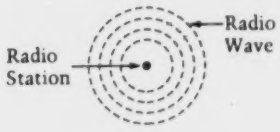
Grade 8 Question: Radio Stations

The Task

This question requires you to show your work and explain your reasoning. You may use drawings, words, and numbers in your explanation. Your answer should be clear enough so that another person could read it and understand your thinking. It is important that you show all your work.

Radio station KMAT in Math City is 200 miles from radio station KGEO in Geometry City. Highway 7, a straight road, connects the two cities.

KMAT broadcasts can be received up to 150 miles in all directions from the station and KGEO broadcasts can be received up to 125 miles in all directions. Radio waves travel from each radio station through the air, as represented below.

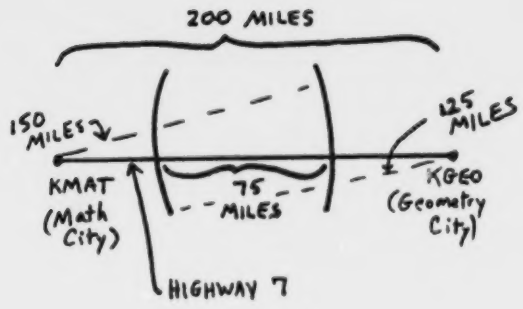


On the next page, draw a diagram that shows the following.

- Highway 7
- The location of the two radio stations
- The part of Highway 7 where both radio stations can be received

Be sure to label the distances along the highway and the length in miles of the part of the highway where both stations can be received.

Possible Solution



There is a 75-mile part of Highway 7 that is within both broadcast areas. It starts 75 miles outside Math City and ends 150 miles outside Math City.

Students need to assimilate and translate semantic information in order to draw a diagram that graphically depicts the location of the radio stations and Highway 7 accurately in terms of given boundary conditions. A graphical approach to this task should enable students to determine the length of the overlapping portion of Highway 7, along which both radio stations can be received. Any satisfactory response must clearly illustrate an overlapping region, whereas, in addition, any extended response must clearly identify the overlap and correctly determine its length to be 75 miles.

National Results, Scoring Guide, and Sample Responses

National Percent for Each Category*	Rating and Performance Category	
16 (1.1)	0	No Response

* The standard errors of the estimated percentages appear in parentheses.

National Percent for Each Category

Rating and Performance Category

45 (1.6)

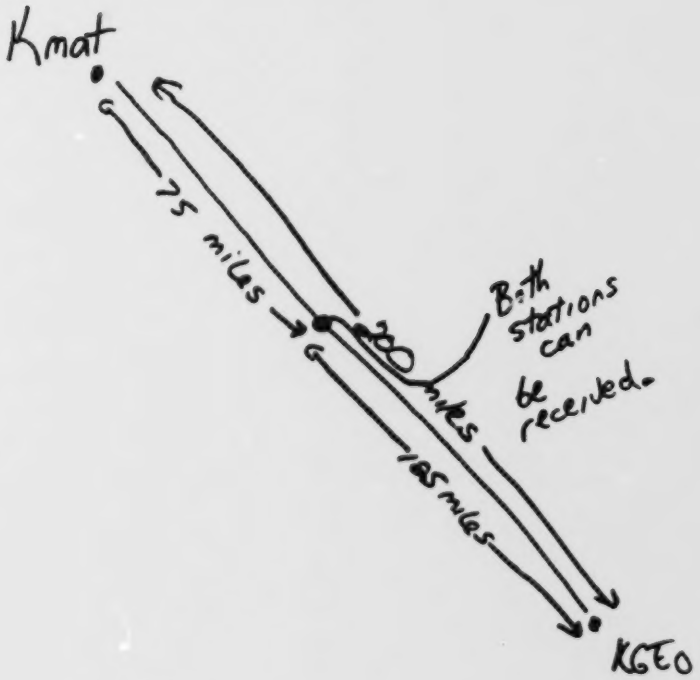
1 **Incorrect -- The work is completely incorrect or irrelevant, or the response states, "I don't know."**

This INCORRECT response does not relate the information given in the problem in a manner that conveys either a meaningful problem solving approach or an adequate solution.

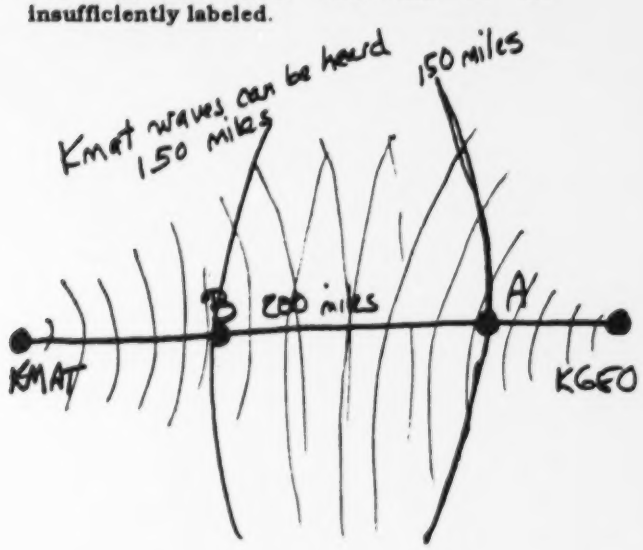
22 (1.2)

2 **Minimal -- Diagram with only cities, Highway 7, and 200 miles labeled; or a diagram that shows some, but not all, of the given distances: 125, 150, or 200 miles. Minimal responses do not recognize that the common broadcast area is a length along the highway.**

This MINIMAL response correctly depicts two pieces of information (radio stations KMAT and KGEO are 200 miles apart and station KGEO can broadcast 125 miles) and shows rudimentary understanding. It does not show the common broadcast area as a length along the highway.



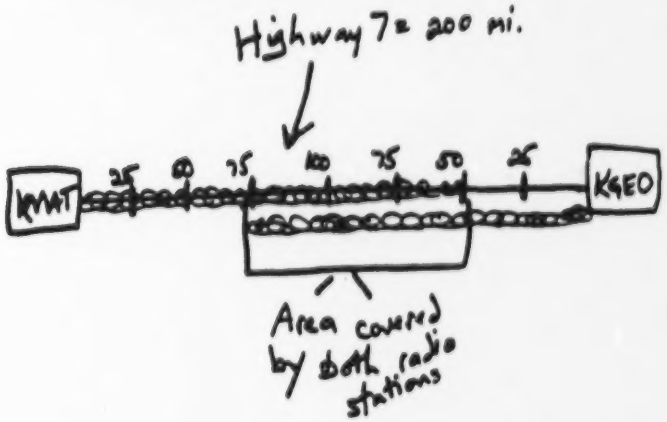
National Percent for Each Category	Rating and Performance Category	
13 (0.9)	3	**Partial -- Diagram with cities, Highway 7, and 200 miles labeled and identification of common broadcast area as a length along (or not on) the highway. Two or more of the radio wave distances 250, 125, and 75 are insufficiently labeled.**



The can both be heard between A & B

This is PARTIAL response indicates considerable understanding of the task relative to the given information. The diagram shows the radio stations to be 200 miles apart and that KMAT can broadcast 150 miles. Additionally, the diagram shows a part of the highway (from A to B) along which both radio stations can be heard. However, the response does not show the broadcast range of station KGEO and does not indicate the length of the common broadcast area.

4 (0.5)	4	**Satisfactory -- Diagram with cities, Highway 7, 200 miles, and all radio wave distances labeled and identification of common broadcast area on Highway 7 as a length. At the same time, omits or incorrectly computes length of the highway along which both radio stations can be received.**



This SATISFACTORY diagram shows a good understanding of the problem. Although the student correctly labeled the common area along Highway 7 where the two stations could be heard, the length in miles of this region was not indicated.

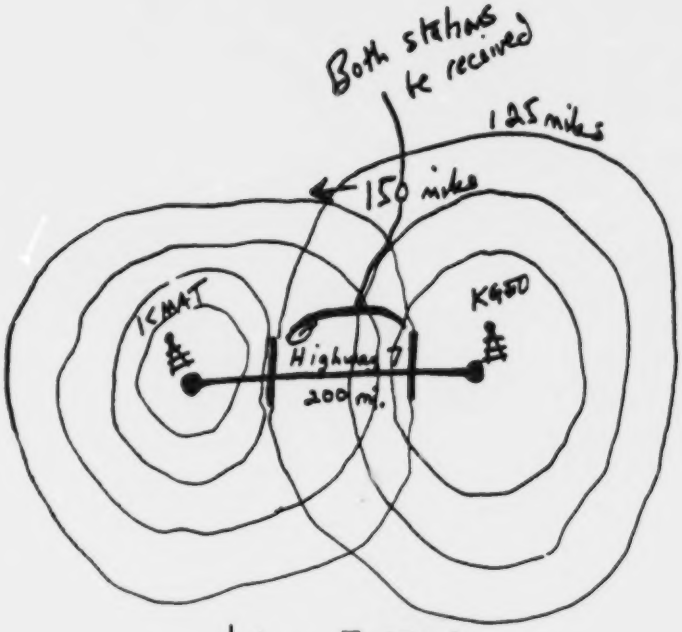
National Percent for Each Category

1 (0.3)

Rating and Performance Category

5 **Extended -- An accurate, well-labeled diagram (as described in the score 4 category) clearly indicating that the portion of Highway 7 along which both radio stations can be received is 75 miles in length.**

This is a solid EXTENDED response. The diagram is accurate and well labeled. Additionally, below the diagram a statement correctly concludes that the length of the part of Highway 7 along which both radio stations can be heard is 75 miles.



Highway 7 200 miles
KMAT 150 mile radius
KGEO 125 mile radius
Part of Highway 7 where both radio stations can be received 75

Performance Highlights: Radio Stations

An important aspect of mathematical power is the need to use logic and diagrams to make sense of a situation and to communicate this reasoning. However, as the results in TABLE 2.9 indicate, many students have yet to recognize that diagrams can be effective analytical and communications tools. Even though a variety of diagrams or explanations could be used to help explain the intersection of the broadcast areas of the two radio stations and no particular approach was preferred, only 5 percent of the eighth graders were able to read and interpret the question and translate this information to develop a labelled model that represented the situation.

TABLE 2.9 National Results for Demographic Subgroups for the Extended-Response Task, "Radio Stations"

Grade 8

	No Response	Incorrect	Minimal	Partial	Satisfactory	Extended	Satisfactory or Better
Nation	16 (1.1)	45 (1.6)	22 (1.2)	13 (0.9)	4 (0.5)	1 (0.3)	5 (0.6)
Northeast	15 (1.7)	42 (3.5)	22 (2.0)	15 (2.5)	5 (1.2)	1 (0.4)	6 (1.1)
Southeast	18 (2.1)	50 (2.8)	17 (1.7)	12 (1.4)	3 (0.8)	0 (0.3)	3 (0.8)
Central	12 (2.2)	43 (2.0)	26 (2.8)	14 (2.4)	5 (1.3)	1 (0.2)	6 (1.3)
West	17 (2.5)	43 (3.5)	23 (2.5)	10 (1.2)	4 (1.1)	2 (0.9)	6 (1.4)
White	11 (1.2)	40 (2.0)	26 (1.6)	16 (1.2)	5 (0.8)	2 (0.4)	7 (0.9)
Black	32 (4.1)	55 (4.2)	8 (2.1)	4 (1.6)	1 (0.6)	0 (0.0)	1 (0.6)
Hispanic	26 (2.5)	58 (2.9)	11 (2.0)	4 (1.3)	1 (0.6)	0 (0.0)	1 (0.6)
Male	17 (1.2)	46 (2.1)	19 (1.8)	13 (1.3)	4 (0.7)	1 (0.3)	4 (0.8)
Female	14 (1.7)	43 (2.0)	24 (1.7)	12 (1.3)	4 (0.8)	2 (0.5)	6 (1.0)
Advantaged Urban	4 (1.7)	32 (3.2)	30 (3.3)	24 (3.1)	7 (1.8)	3 (1.0)	10 (1.5)
Disadvantaged Urban	38 (4.6)	49 (4.7)	7 (1.7)	4 (1.0)	2 (1.5)	0 (0.0)	2 (1.5)
Extreme Rural	15 (4.9)	39 (7.3)	30 (5.6)	14 (4.8)	2 (1.0)	0 (0.5)	2 (1.2)
Other	15 (1.5)	47 (2.2)	21 (1.4)	12 (0.9)	4 (0.7)	1 (0.4)	5 (0.8)
Public	17 (1.2)	45 (1.8)	21 (1.4)	12 (1.1)	4 (0.6)	1 (0.3)	5 (0.7)
Catholic and Other Private	8 (1.6)	42 (2.7)	25 (2.3)	18 (2.2)	6 (1.3)	1 (0.4)	7 (1.4)

The standard errors of the estimated percentages appear in parentheses. It can be said with about 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details). When the proportion of students is either 0 percent or 100 percent, the standard error is inestimable. However, percentages 99.5 percent and greater were rounded to 100 percent and percentages 0.5 percent or less were rounded to 0 percent. Percentages may not total 100 percent due to rounding error.

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment

Students with incorrect responses provided no evidence that they were able to make sense of the problem, often copying a piece of information from the problem or submitting a meaningless drawing (or both). Forty-five percent of the eighth graders nationally provided such responses and another 16 percent did not answer the question at all. Although this meant the majority appeared to be essentially at a loss as to the nature of this task, about one-third did seem to have some understanding of the information presented in relation to the task required. Approximately 22 percent received minimal credit and another 13 percent received partial credit, the difficulty with these responses being an incomplete approach, at best understandable only by those familiar with the problem. These sketchy solutions appeared in spite of directions explicitly telling students what to diagram and to be sure to label the distances and the part of the highway where both stations can be received.

Across the categories of students by region, race/ethnicity, gender, type of community, and type of school, a majority of only one subgroup provided at least minimal responses: advantaged urban students. From 32 percent (advantaged urban) to 58 percent (Hispanic) of the students by subgroup provided meaningless information.

As shown in TABLE 2.10, the percentages of success for public-school eighth graders in the jurisdictions participating in the Trial State Assessment Program were similar to those for the nation. However, in two states, Iowa and Minnesota, at least 10 percent of the students were estimated to have provided satisfactory or better diagrams. For five states, Iowa, Minnesota, Nebraska, New Hampshire, and North Dakota, the majority of the students were estimated to have provided minimal or better responses.

TABLE 2.10 | **Percentages for Responses to Extended-Response Question, "Radio Stations"**

PUBLIC SCHOOLS	Grade 8 - 1992						
	No Response	Incorrect	Minimal	Partial	Satisfactory	Extended	Satisfactory or Better
NATION	17 (1.2)	45 (1.8)	21 (1.4)	12 (1.1)	4 (0.6)	1 (0.3)	5 (0.7)
Northeast	17 (2.0)	40 (4.2)	22 (2.6)	16 (3.1)	4 (1.5)	1 (0.5)	5 (1.6)
Southeast	19 (2.4)	52 (2.7)	15 (1.8)	11 (1.4)	2 (0.7)	0 (0.3)	3 (0.7)
Central	14 (2.6)	43 (2.5)	25 (3.3)	13 (2.9)	4 (1.3)	1 (0.2)	5 (1.4)
West	17 (2.8)	44 (4.0)	23 (2.7)	9 (1.4)	4 (1.2)	3 (1.0)	7 (1.5)
STATES							
Alabama	22 (2.0)	53 (2.4)	14 (1.5)	8 (1.1)	2 (0.6)	1 (0.4)	3 (0.8)
Arizona	16 (1.7)	48 (2.3)	20 (1.8)	13 (1.6)	2 (0.7)	1 (0.4)	3 (0.9)
Arkansas	13 (1.8)	57 (2.6)	18 (1.8)	8 (1.2)	2 (0.6)	1 (0.3)	3 (0.7)
California	19 (1.7)	44 (2.3)	19 (2.1)	13 (2.1)	3 (0.9)	1 (0.5)	4 (1.1)
Colorado	11 (1.2)	43 (2.0)	23 (1.6)	16 (1.7)	5 (0.9)	2 (0.6)	7 (1.1)
Connecticut	10 (1.2)	44 (2.3)	23 (2.0)	16 (1.7)	6 (1.0)	2 (0.4)	7 (1.2)
Delaware	16 (2.0)	51 (2.4)	17 (1.5)	12 (1.8)	3 (0.9)	1 (0.5)	4 (1.0)
Dist. Columbia	30 (2.1)	57 (2.3)	8 (1.0)	3 (1.1)	1 (0.5)	1 (0.5)	2 (0.7)
Florida	22 (1.8)	46 (2.2)	19 (1.9)	10 (1.3)	3 (0.9)	1 (0.3)	4 (0.9)
Georgia	19 (1.7)	49 (2.1)	19 (1.6)	7 (1.1)	4 (0.9)	1 (0.5)	5 (0.9)
Hawaii	23 (1.8)	47 (2.3)	15 (1.6)	11 (1.3)	3 (0.8)	1 (0.4)	3 (0.8)
Idaho	11 (1.1)	46 (2.2)	21 (2.0)	15 (1.2)	5 (1.0)	2 (0.6)	7 (1.2)
Indiana	7 (1.1)	48 (2.6)	25 (2.1)	15 (1.9)	3 (0.7)	1 (0.6)	5 (1.0)
Iowa	6 (0.9)	35 (1.8)	28 (1.9)	21 (2.0)	6 (0.9)	4 (1.1)	10 (1.2)
Kentucky	13 (1.4)	52 (2.3)	22 (1.7)	10 (1.1)	2 (0.5)	1 (0.4)	3 (0.6)
Louisiana	24 (2.2)	54 (2.1)	15 (1.4)	6 (1.0)	1 (0.4)	0 (0.2)	1 (0.5)
Maine	7 (0.9)	46 (2.1)	21 (1.9)	18 (1.8)	6 (0.9)	2 (0.6)	8 (1.1)
Maryland	15 (1.7)	48 (2.1)	20 (1.8)	13 (1.6)	3 (0.9)	1 (0.5)	5 (1.1)
Massachusetts	12 (1.2)	45 (2.8)	22 (2.4)	14 (1.9)	5 (0.9)	2 (0.7)	7 (1.0)
Michigan	15 (1.6)	46 (1.9)	21 (1.9)	12 (1.1)	5 (0.9)	2 (0.6)	7 (1.1)
Minnesota	6 (1.1)	41 (2.3)	20 (1.9)	21 (2.4)	7 (1.0)	4 (1.0)	11 (1.6)
Mississippi	20 (1.9)	57 (2.2)	14 (1.7)	6 (1.2)	2 (0.6)	0 (0.2)	2 (0.6)
Missouri	10 (1.2)	47 (2.1)	23 (1.9)	12 (1.6)	5 (0.9)	2 (0.7)	8 (1.2)
Nebraska	6 (1.4)	44 (2.8)	23 (2.0)	20 (3.0)	6 (1.0)	1 (0.6)	7 (1.1)
New Hampshire	8 (1.1)	42 (2.1)	24 (1.8)	18 (1.5)	5 (1.0)	3 (0.9)	8 (1.3)
New Jersey	12 (1.6)	46 (2.3)	22 (2.0)	14 (1.8)	5 (0.9)	1 (0.5)	6 (1.0)
New Mexico	17 (1.7)	51 (2.2)	19 (1.6)	11 (1.4)	3 (0.7)	1 (0.3)	3 (0.8)
New York	15 (2.1)	44 (2.4)	22 (1.6)	11 (1.5)	5 (0.9)	2 (0.7)	7 (1.2)
North Carolina	12 (1.4)	54 (2.2)	21 (1.8)	10 (1.5)	2 (0.5)	1 (0.4)	3 (0.7)
North Dakota	6 (1.2)	43 (2.5)	25 (2.0)	18 (1.8)	5 (0.9)	3 (1.2)	8 (1.5)
Ohio	12 (1.3)	47 (2.1)	25 (2.1)	12 (1.8)	5 (1.0)	0 (0.2)	5 (1.0)
Oklahoma	10 (1.4)	46 (2.0)	23 (1.6)	16 (1.7)	3 (0.8)	2 (0.8)	5 (1.1)
Pennsylvania	12 (1.2)	44 (2.1)	22 (1.7)	14 (1.4)	6 (1.1)	3 (0.7)	8 (1.5)
Rhode Island	10 (1.1)	49 (3.5)	25 (2.9)	12 (1.6)	2 (0.8)	2 (0.6)	5 (1.0)
South Carolina	12 (1.4)	55 (2.1)	21 (1.7)	10 (1.2)	2 (0.6)	1 (0.4)	3 (0.6)
Tennessee	16 (1.5)	51 (2.4)	22 (1.7)	10 (1.6)	1 (0.5)	0 (0.2)	2 (0.6)
Texas	16 (1.9)	45 (2.2)	21 (1.5)	12 (1.4)	3 (0.7)	3 (1.1)	5 (1.3)
Utah	10 (1.4)	46 (2.5)	20 (1.7)	16 (1.5)	7 (1.2)	2 (0.4)	8 (1.3)
Virginia	13 (1.5)	48 (2.5)	22 (1.7)	12 (1.4)	4 (0.8)	1 (0.5)	5 (1.0)
West Virginia	13 (1.5)	53 (2.4)	21 (1.7)	10 (1.2)	3 (0.6)	0 (0.3)	3 (0.6)
Wisconsin	9 (1.9)	43 (1.8)	23 (1.8)	18 (1.9)	4 (1.0)	4 (1.0)	8 (1.2)
Wyoming	8 (1.2)	45 (2.4)	22 (1.7)	18 (1.6)	6 (0.9)	3 (0.7)	8 (1.1)
TERRITORIES							
Guam	45 (2.4)	40 (2.4)	10 (1.9)	3 (1.0)	2 (0.8)	1 (0.5)	3 (0.8)
Virgin Islands	47 (3.4)	47 (3.3)	4 (1.1)	1 (0.7)	0 (0.3)	0 (0.0)	0 (0.3)

The standard errors of the estimated percentages appear in parentheses. It can be said with 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details). Percentages may not total 100 percent due to rounding error.

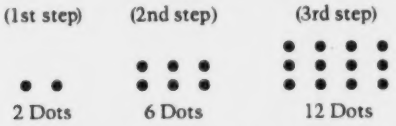
SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment.

Grade 8 Question: Marcy's Dot Pattern

The Task

This question requires you to show your work and explain your reasoning. You may use drawings, words, and numbers in your explanation. Your answer should be clear enough so that another person could read it and understand your thinking. It is important that you show all your work.

A pattern of dots is shown below. At each step, more dots are added to the pattern. The number of dots added at each step is more than the number added in the previous step. The pattern continues infinitely.



Marcy has to determine the number of dots in the 20th step, but she does not want to draw all 20 pictures and then count the dots.

Explain or show how she could do this and give the answer that Marcy should get for the number of dots.

Did you use the calculator on this question?

Yes No

Possible Solution

The explanation should include one of the following ideas with no false statements:

a) For each successive step, the number of rows and the number of columns is increasing by 1, forming a pattern. For example, the first step shows a pattern of dots that consists of one row by two columns, the second step shows a pattern of dots that consists of two rows by three columns, the third step three rows by four columns, and so on. Continuing in this pattern, the twentieth step would have 20 x 21, or 420, dots.

Possible Solution (continued)

b) Look at successive differences between consecutive steps. The differences 4, 6, 8, 10, ... form a pattern. There are 19 differences forming the pattern 4, 6, 8, 10, ... 38, 40, and this sum is equal to (9 x 44) + 22, or 418. However, 2 must be added for the first step, yielding a response of 420.

The solution to this task requires students to analyze several steps in a pattern of dots in order to conjecture about a general rule for determining the number of dots for any particular step in the pattern. Additionally, students are required to use their rule to find the number of dots at a particular step in an extension of the pattern where it no longer is convenient to draw all of the intermediate dot figures. One approach is to think of the steps in the pattern as consisting of dots in rows and columns and to realize that the number of dots in the nth step can be expressed as
$x_n = n(n+1)$ for n= 1,2,3 ... and thus
$x_{20} = 20(20+1) = (20)21 = 420$.

Other approaches are possible and students could use arithmetic or algebraic concepts to explain their reasoning. Although a few students did write an algebraic equation to express a rule for the general term in a recursive relationship, it was neither expected nor necessary for students to do so.

National Results, Scoring Guide, and Sample Responses

National Percent for Each Category*

16 (1.0)

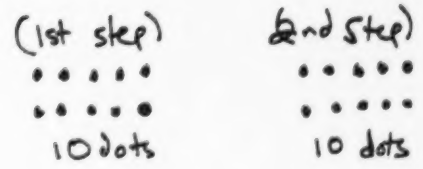
63 (1.3)

Rating and Performance Category

0 **No Response**

1 **Incorrect -- The work is completely incorrect or irrelevant, or the response states, "I don't know."**

It is difficult to discern an explanation for this INCORRECT response. One possibility is that the student apportioned the total of 20 dots in the three steps shown into two 2 x 5 sets.



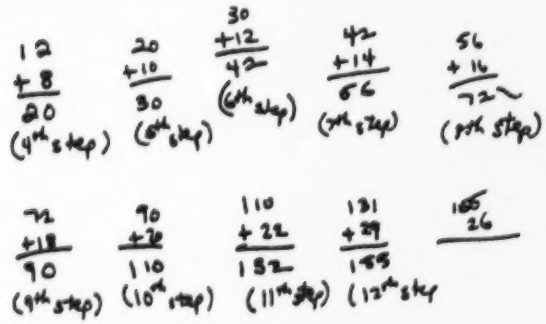
National Percent for Each Category

10 (0.7)

Rating and Performance Category

2 **Minimal -- An attempt to generalize the pattern on a superficial level or to draw all 20 pictures in the pattern (with a clear understanding of the pattern).**

This MINIMAL response illustrates a student's attempt to display the first 12 steps in the pattern. There is some understanding of the number of total dots in each entry but no attempt is made to explain the pattern in terms of rows and columns.



* The standard errors of the estimated percentages appear in parentheses.

National Percent for Each Category*

Rating and Performance Category

6 (0.7)

3 **Partial -- The response has communicated a partially correct generalization of the pattern.**

When the pattern starts with 2 dots the next step is to add 4 dots to it and the 3rd step is to add 6 dots to it so every time there is a new step you add 2 dots to the last amounts you added on to the last step. You would multiply two dots on and on until you reached the 20th step

This PARTIAL response does begin to formulate an explanation of the total number of dots for an entry. However, the last sentence incorrectly uses the term "multiply" in an attempt to discuss the 20th step. At this point, the explanation falters.

National Percent for Each Category

Rating and Performance Category

1 (0.2)

4 **Satisfactory -- The response contains a completely correct generalization of the pattern but does not include -- or incorrectly states -- the number of dots (420) in the 20th step.**

Multiply each step by 1st higher such as

1 = 1 × 2
2 = 2 × 3
3 = 3 × 4
4 = 4 × 5

This SATISFACTORY response provides sufficient evidence of how to generate the various steps in the pattern by multiplying the number of rows times the number of columns. However, the student does not determine the number of dots in the 20th step.

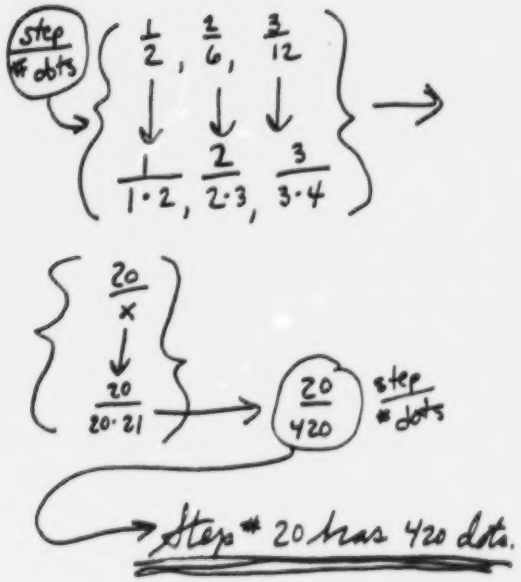
National Percent for Each Category*

5 (0.6)

Rating and Performance Category

5 **Extended -- This response contains a completely correct generalization of the pattern and specifies that there are 420 dots in the 20th step.**

This is a strong EXTENDED response. The student clearly related the number of dots in a step to an appropriate multiplication rule. This student then moves directly from step three to step 20 and determines the correct number of dots for that step.



Performance Highlights: Marcy's Dot Pattern

Because only about half the eighth graders are enrolled in prealgebra (28 percent) or algebra (20 percent),[13] it was expected that this multistep prealgebra question would be difficult for students, given its challenging requirements. First, students needed to identify a pattern in the dots that continues indefinitely. Second, they needed to develop and communicate a generalization to describe any term in the pattern. Finally, they had to apply their generalization to the 20th term in the pattern.

Because the answer could have been given as an algebraic expression for the nth term of an equation, this question was designed to provide the eighth graders enrolled in prealgebra and algebra an opportunity to demonstrate their understanding, while still having several numbers and operations based solutions available for all students. A satisfactory or better response, however, did need to state an accurate recursion rule or provide some computational information with an explanation.

As shown by the results presented in TABLE 2.11, most of the eighth graders fell short of detecting and communicating a pattern in the dots. Nationally, 16 percent did not respond and 63 percent provided irrelevant or inaccurate information about the dot pattern. Of the students who were able to detect the pattern, most were not able to complete the step of generalizing to a rule that could be used to find the dots in any term. Ten percent were able to demonstrate their understanding of the pattern by working some number of the terms beyond those given and another 6 percent tried to provide a generalization. Only 6 percent of the students provided satisfactory or better responses. The 1 percent of the students providing satisfactory rather than extended responses provided a generalization but failed to apply it to the 20th term. The remaining 5 percent worked the problem in full.

In general, the results were relatively consistent across demographic subgroups. However, 13 percent of the advantaged urban students provided satisfactory or better responses compared to 1 percent of the disadvantaged urban students. Also, 13 percent of the private-school eighth graders provided satisfactory or better responses compared to 5 percent of the public-school students.

---

[13] *Data Compendium for the NAEP 1992 Mathematics Assessment of the Nation and the States* (Washington, DC: National Center for Education Statistics, 1993).

TABLE 2.11 National Results for Demographic Subgroups for the Extended-Response Task, "Marcy's Dot Pattern"

Grade 8

	No Response	Incorrect	Minimal	Partial	Satisfactory	Extended	Satisfactory or Better
Nation	16 (1.0)	63 (1.3)	10 (0.7)	6 (0.7)	1 (0.2)	5 (0.6)	6 (0.7)
Northeast	18 (3.2)	61 (3.2)	10 (1.9)	4 (0.7)	2 (0.5)	6 (1.8)	8 (1.6)
Southeast	20 (2.0)	64 (2.2)	9 (1.5)	3 (0.7)	1 (0.4)	4 (1.1)	4 (1.3)
Central	10 (1.5)	65 (2.1)	10 (1.4)	8 (1.4)	1 (0.4)	6 (1.1)	7 (1.4)
West	16 (2.0)	62 (2.8)	10 (1.1)	7 (1.8)	0 (0.2)	4 (1.1)	4 (1.1)
White	12 (1.1)	63 (1.5)	11 (0.8)	7 (0.8)	1 (0.2)	6 (0.8)	8 (0.9)
Black	24 (2.9)	67 (2.9)	6 (1.6)	2 (0.9)	0 (0.0)	1 (0.5)	1 (0.5)
Hispanic	28 (2.8)	61 (3.1)	7 (2.0)	3 (1.2)	0 (0.0)	1 (0.5)	1 (0.5)
Male	19 (1.5)	63 (2.2)	8 (1.0)	5 (0.9)	1 (0.2)	5 (0.9)	5 (0.9)
Female	13 (1.2)	63 (1.6)	12 (1.1)	6 (1.0)	1 (0.3)	5 (0.8)	6 (0.9)
Advantaged Urban	8 (2.9)	62 (5.1)	10 (1.9)	6 (1.6)	1 (0.6)	11 (2.5)	13 (2.6)
Disadvantaged Urban	32 (3.9)	59 (4.7)	4 (1.3)	4 (1.9)	1 (0.6)	1 (0.5)	1 (0.7)
Extreme Rural	16 (2.9)	69 (3.6)	8 (2.3)	2 (1.1)	1 (0.7)	4 (2.0)	5 (2.3)
Other	15 (1.3)	62 (1.5)	11 (0.9)	6 (0.9)	1 (0.2)	4 (0.7)	5 (0.7)
Public	16 (1.2)	64 (1.4)	9 (0.8)	6 (0.7)	1 (0.2)	4 (0.6)	5 (0.6)
Catholic and Other Private	11 (1.7)	56 (2.7)	12 (1.6)	7 (1.2)	2 (0.9)	10 (2.2)	13 (2.0)

The standard errors of the estimated percentages appear in parentheses. It can be said with about 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details). When the proportion of students is either 0 percent or 100 percent, the standard error is inestimable. However, percentages 99.5 percent and greater were rounded to 100 percent and percentages 0.5 percent or less were rounded to 0 percent. Percentages may not total 100 percent due to rounding error.

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment

The results for participating states and jurisdictions, shown in TABLE 2.12, mirror those for the nation. The reasoning and analysis required of students to conjecture, describe, and use a general approach for extending the pattern of dots proved to be a considerable challenge for most students. In general, 80 to 90 percent of the eighth graders attempted this task, but about 60 to 70 percent of them could not find or articulate the pattern in the dots. However, in Connecticut, Maine, Minnesota, New Jersey, and Wisconsin, an estimated 30 percent or more of the students identified a pattern, receiving minimal credit or better for their responses. In four states, an estimated 10 percent or more of the eighth graders provided responses judged satisfactory or better -- Connecticut, Maine, New Jersey, and Wisconsin.

TABLE 2.12 | **Percentages for Responses to Extended-Response Question, "Marcy's Dot Pattern"**

PUBLIC SCHOOLS	Grade 8 - 1992						
	No Response	Incorrect	Minimal	Partial	Satisfactory	Extended	Satisfactory or Better
NATION	16 (1.2)	64 (1.4)	9 (0.8)	6 (0.7)	1 (0.2)	4 (0.6)	5 (0.6)
Northeast	18 (3.9)	62 (3.8)	9 (1.9)	4 (1.0)	1 (0.5)	6 (1.6)	7 (1.8)
Southeast	22 (2.2)	63 (2.3)	9 (1.5)	3 (0.7)	1 (0.6)	2 (0.9)	3 (1.1)
Central	10 (1.3)	66 (2.3)	9 (1.6)	8 (1.5)	1 (0.4)	6 (1.2)	7 (1.4)
West	17 (2.2)	63 (2.8)	10 (1.3)	7 (1.9)	0 (0.3)	3 (1.1)	4 (1.1)
STATES							
Alabama	15 (1.5)	70 (2.2)	8 (1.2)	4 (0.8)	0 (0.2)	3 (0.7)	3 (0.7)
Arizona	17 (1.4)	62 (2.3)	9 (1.2)	6 (1.1)	1 (0.5)	4 (0.9)	6 (1.1)
Arkansas	14 (1.8)	70 (2.3)	9 (1.5)	3 (0.9)	2 (0.4)	3 (0.7)	4 (0.9)
California	20 (1.6)	56 (2.5)	11 (1.6)	6 (1.0)	1 (0.5)	5 (0.9)	6 (1.1)
Colorado	12 (1.4)	61 (1.8)	13 (1.1)	8 (1.1)	2 (0.6)	5 (0.8)	6 (1.1)
Connecticut	11 (1.3)	58 (2.0)	12 (1.5)	9 (1.0)	2 (0.7)	8 (1.0)	10 (1.3)
Delaware	17 (2.2)	63 (2.5)	10 (1.4)	5 (1.0)	1 (0.4)	4 (1.1)	5 (1.1)
Dist. Columbia	24 (2.3)	64 (2.4)	6 (1.4)	2 (0.8)	1 (0.5)	4 (0.9)	5 (0.9)
Florida	19 (1.8)	64 (2.5)	8 (1.4)	4 (1.0)	1 (0.3)	4 (0.7)	5 (0.9)
Georgia	17 (1.7)	65 (2.1)	10 (1.4)	3 (0.7)	1 (0.5)	4 (0.9)	5 (0.9)
Hawaii	24 (1.7)	56 (2.2)	11 (1.5)	6 (0.9)	2 (0.6)	3 (0.7)	4 (0.9)
Idaho	13 (1.5)	63 (1.8)	9 (1.5)	9 (1.3)	1 (0.4)	5 (1.1)	6 (1.3)
Indiana	12 (1.1)	65 (2.1)	11 (1.6)	6 (0.9)	1 (0.5)	5 (0.9)	6 (1.0)
Iowa	6 (0.8)	67 (2.2)	11 (1.6)	8 (1.0)	3 (0.6)	6 (0.9)	8 (1.1)
Kentucky	10 (1.4)	68 (1.9)	11 (1.1)	5 (0.8)	1 (0.4)	4 (1.0)	5 (1.1)
Louisiana	18 (1.8)	66 (2.3)	9 (1.4)	3 (0.9)	1 (0.5)	1 (0.5)	3 (0.8)
Maine	9 (1.1)	60 (2.4)	15 (1.7)	6 (1.4)	2 (0.7)	7 (1.3)	10 (1.5)
Maryland	15 (2.0)	61 (2.4)	9 (1.2)	6 (1.1)	3 (0.9)	6 (1.1)	9 (1.4)
Massachusetts	14 (1.5)	57 (2.4)	13 (1.6)	7 (1.3)	2 (0.7)	7 (1.1)	9 (1.4)
Michigan	14 (1.6)	64 (2.0)	10 (1.3)	5 (0.9)	2 (0.5)	5 (1.2)	7 (1.1)
Minnesota	7 (1.0)	60 (2.0)	15 (1.6)	10 (1.2)	2 (0.5)	7 (1.1)	9 (1.2)
Mississippi	17 (2.1)	70 (2.2)	7 (1.0)	3 (0.6)	1 (0.5)	2 (0.7)	3 (0.9)
Missouri	11 (1.5)	64 (2.2)	11 (1.5)	6 (1.1)	1 (0.4)	7 (1.2)	8 (1.2)
Nebraska	9 (1.1)	64 (2.2)	12 (1.6)	6 (1.5)	2 (0.6)	7 (1.4)	9 (1.4)
New Hampshire	12 (1.5)	59 (2.3)	12 (1.5)	9 (1.3)	2 (0.6)	7 (1.1)	9 (1.2)
New Jersey	12 (1.6)	58 (2.5)	14 (1.6)	7 (1.3)	2 (0.8)	7 (1.3)	10 (1.6)
New Mexico	15 (1.5)	64 (2.1)	11 (1.5)	5 (1.1)	1 (0.4)	3 (0.7)	4 (0.8)
New York	16 (1.7)	61 (2.1)	11 (1.7)	7 (1.4)	1 (0.4)	4 (1.0)	5 (1.0)
North Carolina	14 (1.6)	66 (1.7)	9 (1.2)	5 (1.0)	1 (0.4)	5 (0.9)	6 (0.9)
North Dakota	8 (1.3)	66 (2.8)	12 (1.6)	7 (1.4)	2 (0.5)	5 (1.0)	7 (1.1)
Ohio	14 (1.4)	60 (2.6)	14 (2.2)	5 (0.8)	1 (0.5)	5 (1.1)	6 (1.1)
Oklahoma	11 (1.4)	67 (2.7)	11 (1.8)	4 (1.0)	1 (0.3)	5 (1.3)	6 (1.3)
Pennsylvania	12 (1.5)	63 (2.2)	12 (1.4)	6 (1.0)	1 (0.5)	6 (1.1)	7 (1.3)
Rhode Island	13 (1.5)	62 (2.3)	11 (2.0)	7 (1.6)	2 (0.6)	5 (1.1)	7 (1.2)
South Carolina	12 (1.4)	68 (1.9)	9 (1.3)	6 (1.0)	1 (0.4)	5 (0.8)	5 (1.0)
Tennessee	15 (1.4)	66 (2.2)	11 (1.9)	4 (0.8)	0 (0.3)	3 (0.7)	3 (0.8)
Texas	16 (1.6)	61 (2.4)	11 (1.4)	7 (1.2)	1 (0.4)	4 (0.9)	6 (1.0)
Utah	12 (1.5)	65 (1.8)	9 (1.1)	8 (1.2)	2 (0.6)	5 (0.8)	6 (1.0)
Virginia	14 (1.5)	62 (1.9)	12 (1.5)	6 (0.9)	2 (0.5)	5 (0.9)	7 (1.1)
West Virginia	16 (1.6)	68 (2.1)	8 (1.0)	4 (0.8)	1 (0.4)	3 (0.7)	4 (0.8)
Wisconsin	9 (1.5)	58 (2.4)	14 (1.7)	7 (1.2)	4 (1.6)	7 (1.2)	11 (2.0)
Wyoming	11 (1.4)	62 (2.1)	13 (1.5)	7 (1.1)	1 (0.5)	6 (1.1)	7 (1.1)
TERRITORIES							
Guam	33 (2.8)	52 (2.8)	7 (1.5)	6 (1.4)	2 (0.8)	0 (0.3)	2 (0.8)
Virgin Islands	48 (2.8)	49 (2.8)	3 (1.0)	0 (0.3)	0 (0.0)	0 (0.0)	0 (0.0)

The standard errors of the estimated percentages appear in parentheses. It can be said with 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details). Percentages may not total 100 percent due to rounding error.

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment.

BEST COPY AVAILABLE

Extended-Response Questions: Grade 12

High-school seniors were administered six extended-response tasks as part of NAEP's 1992 national assessment. (Again, the Trial State Assessment Program was administered at grades 4 and 8, but not at grade 12.) The three tasks released to the public are discussed in this section.

The "Effective Tax Rates" question asks students to work with a definition concerning a tax rate. Students could have approached the problem using either numbers and operations or algebra and functions to represent and model a situation involving variables. The first part of the problem, requiring application of the understanding of an effective tax rate of 5 percent, should be accessible to most high-school seniors. However, the second part -- determining whether or not the tax rate could be 6 percent -- requires an understanding of the conditions under which an equation has no real solution or some understanding of the concept of limit. Students were provided with a scientific calculator for this question.

In "Patterns of Squares," students were asked to use elementary algebraic concepts as well as basic numbers facts to explain why a statement is always true about the relationship among the squares of positive integers that end in the digit 5. That is, explain why, when positive integers ending in the digit 5 are squared, the resulting integer always ends in 25. Twelfth graders also were provided with a scientific calculator to use in answering this question.

The question about "Graphing the Path of an Object" required students to apply geometric and algebraic concepts usually encountered in college preparatory mathematics courses. More specifically, students needed working familiarity with the Pythagorean relationship, the rectangular coordinate system, the concept of slope, and graphic models. They were asked to graph the path of an object and answer questions based on their graph.

Grade 12 Question: Effective Tax Rates

The Task

This question requires you to show your work and explain your reasoning. You may use drawings, words, and numbers in your explanation. Your answer should be clear enough so that another person could read it and understand your thinking. It is important that you show all your work.

One plan for a state income tax requires those persons with income of $10,000 or less to pay no tax and those persons with income greater than $10,000 to pay a tax of 6 percent only on the part of their income that exceeds $10,000.

A person's effective tax rate is defined as the percent of total income that is paid in tax.

Based on this definition, could any person's effective tax rate be 5 percent? Could it be 6 percent? Explain your answer. Include examples if necessary to justify your conclusions.

Did you use the calculator on this question?

Yes No

Possible Solution

a) Yes, it can be 5%: Let x equal the number of dollars of income.

$0.06(x - 10{,}000) = 0.05x$
$0.06x - 600 = 0.05x$
$0.01x = 600$
$x = 60{,}000$

If the income is $60,000 then the **effective** tax rate is 5 percent.

Possible Solution (continued)

b) No, it cannot be 6 %.

Let x equal the number of dollars of income.

$0.06 = 0.06(x - 10{,}000)$
$0.06x = 0.06x - 600$
$0 = -600$; This is a false statement, therefore there is no amount of income for which the effective tax rate is 6 percent;

OR, for an income of x dollars, where x is greater than \$10,000, the amount of tax equals $0.06(x - \$10{,}000)$.

The effective tax rate is:

$$\frac{0.06\ (x\ -\ 10{,}000)}{x} = \frac{0.06x\ -\ 600}{x} = 0.06\ -\ \frac{600}{x}$$

As x becomes very large, the **effective** tax rate approaches 0.06 but theoretically never becomes 6 percent;

OR, $\frac{600}{x}$ is always a positive number, so $0.06 - \frac{600}{x}$ is always less than 0.06.

Thus, the **effective** tax rate is less than 6 percent.
Students need to understand that in order for a person to pay any state tax, his or her income must exceed \$10,000. Thus, an appropriate strategy for this problem would be to represent the amount of a person's taxable income in a meaningful way with the use of an expression such as $(x - 10{,}000)$, where x is the number of dollars of income and x is greater than 10,000. Students then can determine by either arithmetic or algebraic methods that there is a unique income, \$60,000, for which the **effective** tax rate is 5 percent.

In attempting to determine whether there is an income for which the **effective** tax rate is 6 percent, it is necessary for students to extend their reasoning skills to consider either implicitly or explicitly a limiting process or to understand the conditions for which an equation has no real solutions.

National Results, Scoring Guide, and Sample Responses

National Percent for Each Category*	Rating and Performance Category	
20 (1.2)	0	No Response
66 (1.4)	1	Incorrect -- The work is completely incorrect or irrelevant, or the response states, "I don't know."

This INCORRECT response does reiterate some given information but does not display any evidence of an approach that might be used to determine the possibility of either a 5 or 6 percent effective tax rate.

1 plan

income $10,000 or less => no tax

income > $10,000 => pay 6% on difference from $10,000

Any person's effective tax rate cannot be 6-percent or be 5 percent.

If everyone does not pay tax because of their income then any person's effective income could not be.

* The standard errors of the estimated percentages appear in parentheses.

National Percent for Each Category	Rating and Performance Category	
9 (0.9)	2	**Minimal -- Student shows some evidence of working with the 5% or 6% and the $10,000 appropriately.**

This MINIMAL response shows that the student understands that only income over $10,000 is taxed. This is illustrated by the example involving $20,000.

- Yes, it could be 5% if they made enough money to pay 5% of their total salery.

- Yes, it could be 6%, for the same reason as above.

If you made more than $10,000, you pay tax on what is over 10,000, for example, if you made 20,000, you would pay $600 on the extra $10,000 you made.

National Percent for Each Category

2 (0.4)

This PARTIAL response contains a correct illustration of an effective tax rate of 5.4%, which would occur for an income of $100,000. The discussion about an effective tax rate of 6% is inaccurate.

Rating and Performance Category

3 **Partial -- There is evidence of some correct work; i.e., an example of a specific effective tax rate or a relevant equation is displayed.**

$10,000 – tax free

– ($100,000 × .05 = $5,000) ⇒
Effective Tax

100,000 – 10,000 = 90,000 × .06 =
5400

5,400 is what % of 100,000

$\frac{5400}{100000} = 5.4\%$

If your salary is 100,000 it is possible to be taxed 5% of your total income, even if the 1st 10,000 is tax free

5% tax

Income = 200,000 × (.05) = 10,000 ⇒
Effective tax rate

190,000 × (.06) = 11,400 ⇒
taxes paid

6% is possible

National Percent For Each Category	Rating and Performance Category	
2 (0.5)	4	**Satisfactory -- Student correctly shows that the effective tax rate can be 5% OR shows that an effective tax rate of 6% is not possible -- but not both.**

This SATISFACTORY response shows that the amount of income for which the effective rate is 5% must be $60,000.

Let's say income =

$60,000

−10,000

50,000

Tax = .06(50,000) = 3,000

60,000X = 3,000

X = .05 = 5% It could be 5%

I don't see how it could be 6%

National Percent for Each Category

1 (0.4)

This EXTENDED response provides all the work necessary to show that there exists an income for which the effective tax rate is 5%. In a similarly efficient manner, the student demonstrates that there exists no amount of income for which the effective tax rate is 6%.

Rating and Performance Category

5 Extended -- The work for both the 5% and 6% effective tax rate cases is clearly and accurately shown.

YES

5% $\leq \$10{,}000$ no tax

$> \$10{,}000$ 6% on over $10,000

X = $ income over $10,000

$.06(x) = .05(10{,}000 + x)$

$.06x = 500 + .05x$

$.01x = 500$

$x = 50{,}000$

Someone with an income of $60,000 would have effective tax rate of 5%

NO

6%

$.06(x) = .06(10{,}000 + x)$

$.06x = 600 + .06x$

$0 \neq 600$

Not possible to have 6% effective tax rate

Performance Highlights: Effective Tax Rates

TABLE 2.13 contains the national results for the Effective Tax Rates question. One-fifth of the high-school seniors left their papers blank. Nearly two-thirds (66 percent) did not demonstrate understanding of the initial premise of the question -- that only income in excess of $10,000 would be taxed. In most cases, these students simply reiterated some information from the problem situation without making any progress toward answering either of the questions.

TABLE 2.13 National Results for Demographic Subgroups for the Extended-Response Task, "Effective Tax Rates"

Grade 12

	No Response	Incorrect	Minimal	Partial	Satisfactory	Extended	Satisfactory or Better
Nation	20 (1.2)	66 (1.4)	9 (0.9)	2 (0.4)	2 (0.5)	1 (0.4)	3 (0.7)
Northeast	21 (2.5)	62 (3.0)	9 (1.9)	2 (0.7)	4 (1.7)	1 (0.5)	5 (1.8)
Southeast	23 (2.2)	68 (2.7)	7 (1.5)	2 (0.6)	1 (0.4)	0 (0.1)	1 (0.4)
Central	15 (2.2)	69 (2.5)	9 (1.4)	2 (0.6)	2 (0.8)	2 (1.2)	4 (1.7)
West	22 (2.8)	62 (3.4)	10 (1.9)	3 (0.9)	2 (0.9)	1 (0.5)	3 (1.2)
White	17 (1.3)	66 (1.6)	10 (1.1)	3 (0.5)	3 (0.6)	2 (0.5)	4 (0.9)
Black	31 (3.3)	66 (3.7)	1 (0.8)	1 (0.5)	2 (1.7)	0 (0.0)	2 (1.7)
Hispanic	32 (4.7)	63 (6.4)	5 (2.7)	0 (0.3)	0 (0.0)	0 (0.4)	4 (0.4)
Male	21 (1.9)	67 (2.0)	7 (1.0)	2 (0.4)	2 (0.6)	1 (0.4)	3 (0.8)
Female	20 (1.4)	64 (1.7)	10 (1.2)	2 (0.6)	2 (0.8)	1 (0.7)	4 (1.1)
Advantaged Urban	11 (2.0)	64 (3.1)	13 (2.4)	3 (1.0)	5 (1.9)	4 (1.2)	9 (2.1)
Disadvantaged Urban	30 (4.0)	65 (4.3)	3 (1.4)	0 (0.4)	1 (0.8)	0 (0.3)	1 (0.8)
Extreme Rural	16 (2.2)	68 (2.9)	12 (2.8)	3 (1.0)	0 (0.3)	0 (0.4)	1 (0.4)
Other	21 (1.6)	65 (1.9)	8 (1.0)	2 (0.5)	2 (0.6)	1 (0.5)	3 (0.9)
Public	21 (1.3)	66 (1.6)	9 (1.0)	2 (0.4)	2 (0.5)	1 (0.4)	2 (0.6)
Catholic and Other Private	14 (2.1)	63 (2.6)	10 (2.6)	3 (1.0)	6 (1.8)	3 (1.1)	9 (2.1)

The standard errors of the estimated percentages appear in parentheses. It can be said with about 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details). When the proportion of students is either 0 percent or 100 percent, the standard error is inestimable. However, percentages 99.5 percent and greater were rounded to 100 percent and percentages 0.5 percent or less were rounded to 0 percent. Percentages may not total 100 percent due to rounding error.

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment

Nine percent of the twelfth-grade students showed an initial grasp of the foundation of the question, taxation of income above $10,000, but went no further. Two percent received partial credit for accurately computing a tax rate of an income above $10,000, even though the rate was not 5 percent. Most students who produced responses judged to be at the minimal or partial level used arithmetic to show some understanding of the problem situation.

The 2 percent providing satisfactory responses found the income with an effective tax rate of 5 percent, but did not solve the second part of the question. Only 1 percent of the students at grade 12 provided explanations for both parts of the question. Almost all students who gave satisfactory or better responses used algebraic approaches together with the idea of a limit.

By and large, these low levels of performance held across subgroups. However, 9 percent of the advantaged urban students -- as well as of the private-school students -- provided satisfactory or better responses to this task.

Grade 12 Question: Patterns of Squares

The Task

This question requires you to show your work and explain your reasoning. You may use drawings, words, and numbers in your explanation. Your answer should be clear enough so that another person could read it and understand your thinking. It is important that you show all your work.

$$15^2 = 225$$
$$25^2 = 625$$
$$35^2 = 1225$$

The examples above suggest the following statement.

> When a positive integer that ends in the digit 5 is squared, the resulting integer ends in 25.

Explain why this statement is always true. (Hint: $(10n + 5)^2 =$?)

Did you use the calculator on this question?

Yes No

Possible Solution

For n a positive integer:
$(10n + 5)^2 = 100n^2 + 100n + 25 = 100 (n^2 + n) + 25$

Since n is a positive integer it follows that n^2, $n^2 + n$, and $100 (n^2 + n)$ are positive integers. The integer $100 (n^2 + n)$ is a multiple of 100 and thus ends in 00, i.e., its unit and tens digits are both 0. Therefore, when 25 is added to $100 (n^2 + n)$ the sum will end in 25, i.e., the tens and unit digits are 2 and 5, respectively.

When asked to square an expression such as (a + b) many students will **incorrectly** state that $(a + b)^2 = a^2 + b^2$. This significant misconception usually occurs because students fail to recognize that $(a + b)^2 = a^2 + ba + ab + b^2$. Therefore, when the multiplication of (a + b) times (a + b) is carried out and like terms are collected, the resulting product is $a^2 + 2ab + b^2$. Thus, it is the middle term, 2ab, that is often overlooked, even by some good students. This principle and resulting algorithm are central to showing that when a positive integer that ends with a units digit of 5 is squared, the resulting product is an integer that ends in 25. Additionally, it is necessary for students to demonstrate a clear understanding of place value and powers of 10 in order to fully justify their explanations.

National Results, Scoring Guide, and Sample Responses

National Percent for Each Category*	Rating and Performance Category	
17 (1.4)	0	No Response
64 (1.7)	1	Incorrect -- The work is completely incorrect or irrelevant, or the response states, "I don't know."

This INCORRECT response fails to demonstrate any meaningful work and contains several errors that convey a misunderstanding of important algebraic concepts that are being assessed.

$(10n+5)^2 = (15n)^2 = 225$

If the answer from the equation ends in 5, then the answer will alway be 25.

* The standard errors of the estimated percentages appear in parentheses

National Percent for Each Category*

16 (1.2)

This MINIMAL response gives an additional numerical example not provided in the question that indicates some understanding of the problem.

Rating and Performance Category

2 **Minimal -- Student provides additional numerical examples only or states $(10n + 5)^2 = 100n^2 + 25$ only.**

Any Positive integer times itself will always have 25 at the end of the answer because if $15^2 = 225$ $15 \times 15 = 225$ it ends in 5 so it will have 25 in the answer

Example $45 \times 45 = 2025$

$5 \times 5 = 25$

thats How you get 25 at the end of each answer

National Percent for Each Category

Rating and Performance Category

1 (0.3)

3 **Partial -- Student states $(10n + 5)^2 = 100n^2 + 25$, and provides a partially correct explanation.**

This PARTIAL response shows the relationship between a multiple of 100 and the addition of 25. However, this response also contains the misconception that $(10n + 5)^2 = 100n^2 + 25$.

Because the square of 5 is 25 and the square of 10n is always equal to n squared times 100

ie:

$(10n+5)^2 = x$

if $n = 4 \rightarrow [10(n)]^2 = [40]^2 =$ 1600

$n^2 = 4^2 =$ 16 $\times 100 = 1600$

$5^2 = 25$

$1600 + 25 = 1625$

1 (0.2)

4 **Satisfactory -- Student states that $(10n + 5)^2 = 100n^2 + 100n + 25$ and mentions zero(s). The explanation ties 25 to a multiple of 10 or 100.**

This SATISFACTORY response correctly shows that $(10n + 5)^2 = 100n^2 + 100n + 25$ and also gives a rather weak statement that relates a multiple of 100 and the addition of 25.

$(10n+5)^2 = 100n^2 + 100n + 25$

100 times any number or n leves two empty spases with only the 25 can fill

National Percent for Each Category

1 (0.4)

Rating and Performance Category

5 **Extended -- Student displays a solution that is mathematically accurate and provides a clear and complete explanation.**

This EXTENDED response correctly expands $(10n + 5)^2$ and gives an explanation that for any number n the expansion will have 0s in the last two digits and thus when 25 is added the resulting integer will end in 25.

$(10N + 5)^2 = 100N^2 + 100N + 25$

~~If you multiply out your first~~

For any Number N, $100N^2 + 100N$, will end in -00, which you must add 25 to. This results in the number ending in 25.

Performance Highlights: Patterns of Squares

As indicated by the results in TABLE 2.14, twelfth-grade students had great difficulty with this problem. Seventeen percent did not even attempt a solution, and 64 percent provided bits of information or unrelated mathematics (usually incorrect). Sixteen percent of the students showed a minimal grasp of the issue, by providing a different numerical example such as $45 \times 45 = 2{,}025$. Some of the other minimal responses incorrectly stated that $(10n + 5)^2$ equals $100n^2 + 25$. The few responses (1 percent) given partial credit were based on the same misconception about the formula, but also contained an explanation related to the idea of adding 25 to a multiple of 100.

TABLE 2.14 National Results for Demographic Subgroups for the Extended-Response Task, "Patterns of Squares"

Grade 12

	No Response	Incorrect	Minimal	Partial	Satisfactory	Extended	Satisfactory or Better
Nation	17 (1.4)	64 (1.7)	16 (1.2)	1 (0.3)	1 (0.2)	1 (0.4)	2 (0.4)
Northeast	19 (2.0)	57 (2.7)	20 (2.5)	0 (0.3)	0 (0.4)	4 (1.3)	4 (1.4)
Southeast	20 (3.7)	64 (4.1)	14 (2.4)	1 (0.3)	1 (0.3)	1 (0.5)	1 (0.6)
Central	13 (1.9)	69 (2.0)	15 (2.0)	1 (0.6)	1 (0.6)	1 (0.4)	2 (0.7)
West	17 (2.9)	64 (4.0)	15 (2.6)	4 (1.0)	0 (0.3)	0 (0.3)	1 (0.6)
White	12 (1.4)	67 (1.9)	17 (1.4)	2 (0.4)	1 (0.3)	2 (0.5)	3 (0.6)
Black	36 (5.1)	56 (5.5)	7 (1.8)	1 (0.3)	0 (0.0)	0 (0.0)	0 (0.0)
Hispanic	25 (4.4)	56 (5.2)	17 (4.4)	2 (1.7)	0 (0.0)	0 (0.0)	0 (0.0)
Male	19 (2.0)	62 (2.3)	15 (1.4)	2 (0.5)	0 (0.2)	2 (0.7)	2 (0.7)
Female	15 (1.6)	65 (2.2)	17 (1.7)	1 (0.5)	1 (0.3)	1 (0.4)	2 (0.5)
Advantaged Urban	8 (2.1)	64 (3.4)	19 (3.2)	3 (1.5)	2 (0.7)	4 (1.9)	6 (2.1)
Disadvantaged Urban	28 (4.9)	57 (5.1)	14 (2.4)	1 (0.5)	0 (0.0)	0 (0.0)	0 (0.0)
Extreme Rural	17 (4.5)	70 (5.9)	12 (3.3)	1 (1.1)	0 (0.0)	0 (0.0)	0 (0.0)
Other	17 (1.5)	63 (1.9)	17 (1.8)	1 (0.4)	1 (0.3)	1 (0.4)	2 (0.5)
Public	17 (1.5)	64 (1.9)	15 (1.4)	1 (0.4)	1 (0.2)	1 (0.4)	2 (0.4)
Catholic and Other Private	14 (2.2)	59 (3.2)	21 (2.4)	2 (0.7)	2 (0.8)	2 (0.8)	4 (1.1)

The standard errors of the estimated percentages appear in parentheses. It can be said with about 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details). When the proportion of students is either 0 percent or 100 percent, the standard error is inestimable. However, percentages 99.5 percent and greater were rounded to 100 percent and percentages 0.5 percent or less were rounded to 0 percent. Percentages may not total 100 percent due to rounding error.

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment

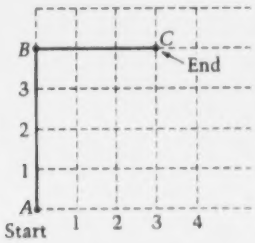
One percent of the responses were judged to be satisfactory and another 1 percent were judged to be extended. This means that an estimated 2 percent of the students accurately conveyed the idea that $(10n + 5)^2 = 100n^2 + 100n + 25$, with only half of them also giving an adequate explanation that related the addition of 25 to a multiple of 100, thereby earning an extended rating. Given that all students do not study algebra, some confusion with the formula may be understandable. Yet, the concepts underlying the question are rooted in a basic understanding of place value combined with multiplication of zeros and fives. It does seem that with some thought, more students would have received partial credit.

Performance across subgroups was quite consistent on this task, because so few students in any group seemed to grasp the idea underlying the question. Even for the advantaged urban students, only 9 percent provided responses judged as partial or better.

Grade 12 Question: Graphing Path of Object

The Task

This question requires you to show your work and explain your reasoning. You may use drawings, words, and numbers in your explanation. Your answer should be clear enough so that another person could read it and understand your thinking. It is important that you show all your work.



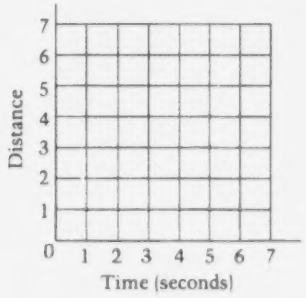
The darkened segments in the figure above show the path of an object that starts at point A and moves to point C at a constant rate of 1 unit per second. The object's distance from point A (or from point C) is the shortest distance between the object and the point.

Please answer the questions on page 9 that refer to this graph.

In the space below, complete the following steps.

a) Sketch the graph of the distance of the object from point A over the 7-second period.

b) Then sketch the graph of the distance of the object from point C over the same period.

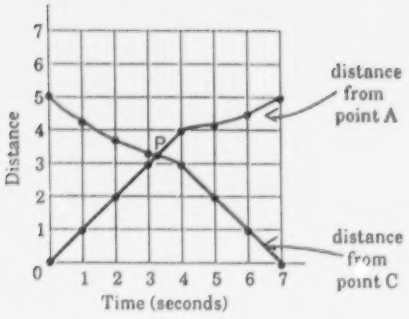


c) On your graph, label point P at the point where the distance of the object from point A is equal to the distance of the object from point C.

d) Between which two consecutive seconds is the object equidistant from points A and C?

Possible Solution

a) and b)



c)

$P = \left(3\frac{1}{8}, \quad 3\frac{1}{8}\right)$

Seconds	Distance from Point A	Distance from Point C
0	0	5
1	1	$\sqrt{18} \approx 4.2$
2	2	$\sqrt{13} \approx 3.6$
3	3	$\sqrt{10} \approx 3.2$
4	4	3
5	$\sqrt{17} \approx 4.1$	2
6	$\sqrt{20} \approx 4.5$	1
7	5	0

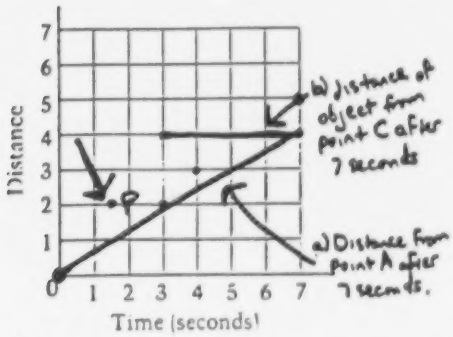
d) Between 3 and 4 seconds.

Students need to realize that the graph of the distance of the object from point A is linear only during the first four seconds. At the end of the fifth second it is critical for students to observe that the distance of the object from point A is equal to the length of the hypotenuse of a right triangle with sides of length 4 and 1 and that distance is equal to $\sqrt{4^2 + 1^2} = \sqrt{17}$ by the Pythagorean relationship. In a like manner, at the end of the sixth and seventh seconds the distance the object is from point A is equal to $\sqrt{4^2 + 2^2} = \sqrt{20} = 2\sqrt{5}$ and $\sqrt{4^2 + 3^2} = \sqrt{25} = 5$, respectively. When the seven resulting (time, distance) ordered pairs are plotted on the axes provided and the graph of the distance of the object from point A is sketched, students should have drawn a non-linear path. The non-linearity may be observed from the change in slope of the path that occurs between the points (4,4) and (5,4.1) and thereafter. The path of the distance of the object from point C, on the other hand, is non-linear for the first four seconds and linear during the final three seconds. Another facet of this task is for students to understand that the distance of the object from point A is equal to the distance of the object from point C at the point where the two curves intersect, which occurs between the third and fourth seconds.

National Results, Scoring Guide, and Sample Responses

National Percent for Each Category*	Rating and Performance Category	
9 (0.8)	0	**No Response**
68 (1.3)	1	**Incorrect -- The work is completely incorrect or irrelevant or the response states, "I don't know."**

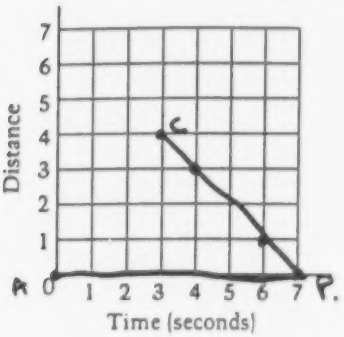
This INCORRECT response indicates some relevance to the task but the work is insufficient to warrant recognition even at the minimal level.



d) 1 and 2

18 (1.4)	2	**Minimal -- At least two points are plotted correctly on at least one of the two distance vs. time graphs.**

This MINIMAL response shows an incomplete graph of the distance of the object from Point C. The portion of the graph shown does contain three correctly plotted points.



d) point C + P

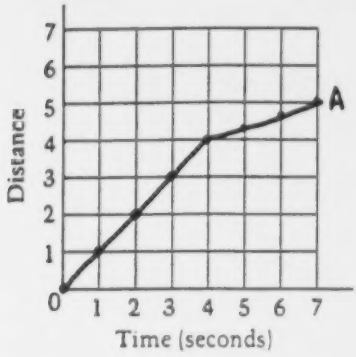
* The standard errors of the estimated percentages appear in parentheses.

National Percent for Each Category

Rating and Performance Category

4 (0.8)

3 **Partial -- Portions of one or both graphs are correct; point P is not located or is located incorrectly and the time when the object is equidistant from points A and C is incorrect or missing.**

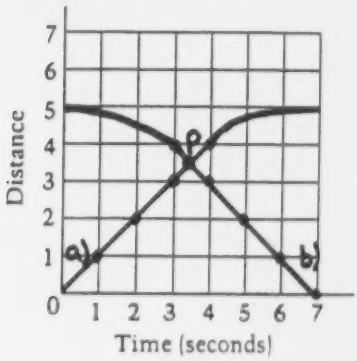


This PARTIAL response correctly shows the graph of the distance of the object from Point A, including the change in slope at the point (4,4) that indicates the path is nonlinear.

d)

0 (0.2)

4 **Satisfactory -- Both graphs are non-linear but the slope of one graph does not change at the appropriate point. Point P is located correctly but the time when the object is equidistant from points A and C is incorrect or missing.**



This SATISFACTORY response would have been at the extended level except that the slope of the graph of the distance of the object from point C does not change at the point (4,3) but rather at the point (3,4). The fact that the curvature of this graph is inaccurate would not have deducted from the student's score.

d) 3 and 4

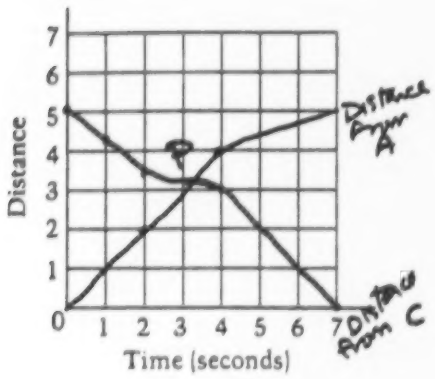
National Percent for Each Category

1 (0.2)

Rating and Performance Category

5 **Extended -- Both graphs are accurately sketched and show change in slope at the points (4,4) and (4,3). Graphs must be non-linear but curvature (concavity) need not be exact. Point P is located at the intersection of the two graphs between the 3rd and 4th seconds.**

This EXTENDED response clearly shows the graphs of the distance of the object from Points A and C are nonlinear curves with the change in the slopes of the curves changing at points (4,4) and (4,3), respectively. Additionally, the student has located point P at the intersection of the two graphs and indicated that the object is equidistant from points A and C between the third and fourth seconds.



d) 3 and 4

Performance Highlights: Graphing Path of Object

One of the most interesting things about the Graphing Path of Object question was the relatively high rate of response. The data presented in TABLE 2.15 show that 91 percent of the twelfth graders tackled this question. However, about two-thirds of the students (68 percent) provided graphs that bore little or no resemblance to the information given in the problem. These findings agree with results from a related constructed-response question administered at grade 12 as part of the California Assessment Program (CAP), which also required graphing and application of the Pythagorean relationship. As stated in the CAP report, "although students have been given opportunities to translate from verbal situations to equations and arithmetic algorithms, they have not had enough experience in proceeding from verbal instructions to geometric figures."[14]

TABLE 2.15 National Results for Demographic Subgroups for the Extended-Response Task, "Graphing Path of Object"

Grade 12

	No Response	Incorrect	Minimal	Partial	Satisfactory	Extended	Satisfactory or Better
Nation	9 (0.8)	68 (1.3)	18 (1.4)	4 (0.8)	0 (0.2)	1 (0.2)	1 (0.3)
Northeast	8 (2.0)	63 (2.8)	20 (3.6)	7 (2.3)	1 (0.3)	1 (0.4)	1 (0.6)
Southeast	10 (1.3)	71 (3.1)	17 (2.9)	1 (0.5)	0 (0.1)	1 (0.4)	1 (0.4)
Central	7 (1.6)	67 (1.7)	20 (2.3)	5 (1.1)	1 (0.5)	1 (0.5)	2 (0.7)
West	10 (1.4)	69 (2.3)	15 (2.0)	5 (1.7)	0 (0.3)	1 (0.5)	1 (0.4)
White	8 (0.9)	66 (1.4)	19 (1.7)	5 (1.1)	1 (0.2)	1 (0.3)	2 (0.4)
Black	12 (2.2)	74 (3.3)	13 (2.6)	1 (0.7)	0 (0.0)	0 (0.0)	0 (0.0)
Hispanic	13 (2.9)	70 (4.2)	16 (3.5)	1 (0.9)	0 (0.0)	0 (0.0)	0 (0.0)
Male	9 (1.2)	64 (1.9)	20 (1.5)	5 (1.1)	1 (0.3)	1 (0.4)	2 (0.5)
Female	8 (1.2)	71 (2.0)	16 (1.9)	4 (0.8)	0 (0.2)	1 (0.3)	1 (0.3)
Advantaged Urban	3 (1.6)	63 (5.2)	24 (4.2)	8 (2.5)	1 (0.4)	1 (0.7)	2 (0.8)
Disadvantaged Urban	20 (2.9)	65 (3.7)	12 (2.2)	4 (1.6)	0 (0.0)	0 (0.0)	0 (0.0)
Extreme Rural	8 (2.7)	68 (4.6)	21 (3.6)	3 (1.3)	0 (0.0)	0 (0.4)	0 (0.4)
Other	8 (0.9)	69 (1.4)	17 (1.7)	4 (1.0)	1 (0.3)	1 (0.3)	2 (0.4)
Public	10 (0.9)	69 (1.4)	17 (1.2)	4 (0.9)	0 (0.1)	1 (0.3)	1 (0.3)
Catholic and Other Private	4 (1.1)	60 (4.2)	27 (5.2)	7 (1.6)	2 (1.0)	1 (0.3)	3 (1.0)

The standard errors of the estimated percentages appear in parentheses. It can be said with about 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details). When the proportion of students is either 0 percent or 100 percent, the standard error is inestimable. However, percentages 99.5 percent and greater were rounded to 100 percent and percentages 0.5 percent or less were rounded to 0 percent. Percentages may not total 100 percent due to rounding error.

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment

[14] *A Question of Thinking: A First Look at Students' Performance on Open-Ended Questions in Mathematics* (Sacramento, CA: California State Department of Education, 1989).

Eighteen percent of the students received a minimal rating, by correctly plotting at least two or three points on the linear portion of one of the time-distance graphs. Four percent of the students were given partial credit for plotting portions of both graphs correctly. Only a very few students, about 1 percent, demonstrated understanding of the Pythagorean relationship in determining the object's shortest distance from either point A or point C. Nearly all students who correctly applied this relationship sketched both graphs correctly, located point P, and determined the time at which the object was equidistant from points A and C. Thus, most students who understood the demands of the task provided extended responses. A very few omitted some piece of the necessary information and received a satisfactory rating.

That only a small percentage of students was able to complete this problem successfully also corresponds to the CAP findings for the question requiring graphing and application of the Pythagorean relationship. As explained in its report, "students' descriptions and diagrams revealed a series of ways they went astray: only 1.5 percent successfully completed the entire problem."

The few performance differences across subgroups were noted for this question at the minimal and partial levels. For example, nearly one-third of advantaged urban students plotted some part, but not all of the information, compared to 16 percent of the disadvantaged urban students. Percentages of satisfactory or better responses were uniformly low across the various subgroups, from 0 to 3 percent.

Summary

The results in this chapter highlight students' difficulties in communicating mathematics ideas and concepts. For some questions, as many as one-fifth of the students left their papers blank, providing no response at all. Often, the majority of students did not provide evidence that they had a grasp of the concepts that needed to be explained, or in some instances that they even understood the question being asked. From approximately one-third to two-thirds of the students provided incorrect responses to the extended questions. Some of this phenomenon could result from students simply not taking enough time to read and understand the question. Or perhaps, students had difficulty in even reading the questions. If students cannot read well, it would influence their ability to do these kinds of mathematics problems.

Some portion of the students did demonstrate understanding of the tasks, but needed more practice in providing complete explanations. In fact, most students who did seem to understand the questions had difficulty explaining their work. Although the percentages of students providing satisfactory or better responses tended to be small, it is encouraging that some students -- from 1 to 16 percent -- provided extended responses to each one of the tasks.

CHAPTER THREE

Summarizing Performance on the Constructed-Response Tasks

In general, there appears to be overwhelming support for assessments that directly address important student learning and foster exemplary teaching practices. In addition to *The NCTM Standards* and Mathematical Sciences Education Board publications such as *Measuring Up*, various national reports, including *Raising Standards for American Education* by the National Council on Education Standards and Testing and *Testing in American Schools: Asking the Right Questions* by the Office of Technology Assessment, recommend such approaches. However, the concerted push toward more instructionally relevant assessment instruments has meant movement away from multiple-choice formats with clearly understood measurement properties to more complicated performance assessment situations. Thus, as information becomes available about various implementations of performance-oriented assessments it is interesting to share the lessons learned.[15]

Difficulty by Question Type and for Subpopulations

TABLES 3.1 and 3.2 summarize performance at grade 4 for the nation and the states on the five extended-response questions included in the 1992 assessment. Across the tasks, three of which were presented in their entirety in Chapter Two, national performance ranged from 10 to 23 percent of the fourth graders providing responses judged as satisfactory or better. On average, 16 percent provided responses judged as satisfactory or better. Across the states, average performance varied from 7 to 22 percent satisfactory or better responses. Seven states had an estimated one-fifth or more of their students provide satisfactory or better responses, on average, including Connecticut, Iowa, Maine, Minnesota, New Hampshire, North Dakota, and Wisconsin.

[15] NAEP shared the results of its pilot study in collecting the nation's writing portfolio in *Exploring New Methods for Collecting Students' School-based Writing* available from the Government Printing Office.

TABLE 3.1 National Percentages of Satisfactory or Better Responses to Extended-Response Questions

Grade 4

	Pizza Comparison	Laura Use Calculator	Graphs of Pockets	Compare Geometric Figures*	Number Patterns (Photo Album)*	Average Percentage Satisfactory or Better
Nation	23 (1.3)	20 (1.5)	10 (0.9)	10 (0.8)	18 (1.1)	16 (0.6)
Northeast	29 (4.3)	25 (3.9)	12 (2.5)	14 (1.6)	25 (3.0)	21 (1.7)
Southeast	20 (2.4)	15 (2.8)	9 (2.3)	10 (1.3)	12 (1.4)	13 (1.3)
Central	23 (2.2)	23 (3.4)	9 (1.7)	8 (1.0)	20 (2.4)	17 (1.0)
West	23 (2.1)	18 (2.3)	10 (1.2)	9 (2.1)	16 (2.2)	15 (1.1)
White	28 (1.7)	24 (2.2)	13 (1.3)	12 (1.1)	22 (1.4)	20 (0.8)
Black	9 (2.1)	5 (1.8)	1 (0.4)	5 (1.3)	4 (1.2)	5 (0.7)
Hispanic	12 (2.8)	11 (2.2)	2 (1.0)	2 (1.0)	8 (2.0)	7 (1.0)
Male	26 (2.0)	22 (1.8)	11 (1.5)	8 (1.2)	14 (1.4)	16 (0.8)
Female	21 (1.4)	18 (2.0)	9 (1.2)	12 (1.3)	22 (1.9)	17 (0.8)
Advantaged Urban	35 (3.6)	32 (4.8)	16 (3.4)	18 (3.4)	28 (4.3)	26 (2.4)
Disadvantaged Urban	14 (3.5)	4 (1.5)	2 (0.9)	4 (1.0)	2 (1.0)	5 (1.0)
Extreme Rural	18 (3.7)	20 (6.3)	10 (2.3)	5 (2.5)	15 (3.1)	14 (1.9)
Other	23 (1.6)	20 (1.7)	10 (1.2)	10 (0.9)	19 (1.4)	17 (0.7)
Public	23 (1.5)	19 (1.6)	10 (1.0)	10 (0.9)	18 (1.3)	16 (0.7)
Catholic and Other Private	23 (2.7)	28 (2.6)	12 (2.0)	12 (1.8)	20 (2.1)	19 (1.1)

*Secure questions, unreleased.

The standard errors of the estimated percentages appear in parentheses. It can be said with about 95 percent certainty for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details).

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment

TABLE 3.2 | **Percentage of Satisfactory or Better Responses to Extended-Response Questions, Grade 4**

PUBLIC SCHOOLS	Grade 4 - 1992					
	Pizza Comparison	Laura's Calculator Correction	Graphs of Pockets	Compare Geometric Figures*	Number Patterns* (Photo Album)	Average Percent Satisfactory or Better
NATION	23 (1.5)	19 (1.6)	10 (1.0)	10 (0.9)	18 (1.3)	16 (0.7)
Northeast	29 (5.1)	24 (4.3)	12 (2.5)	13 (2.2)	26 (3.5)	21 (2.0)
Southeast	20 (2.7)	13 (3.4)	8 (2.4)	10 (1.5)	12 (1.3)	13 (1.6)
Central	23 (2.8)	23 (3.7)	8 (2.1)	8 (1.1)	19 (3.0)	16 (1.3)
West	23 (2.1)	16 (2.2)	10 (1.4)	9 (2.1)	16 (2.3)	15 (1.2)
STATES						
Alabama	16 (1.7)	16 (1.9)	5 (0.9)	6 (0.9)	14 (1.3)	11 (0.8)
Arizona	19 (1.6)	15 (1.2)	7 (1.0)	9 (1.4)	17 (1.3)	13 (0.7)
Arkansas	20 (1.7)	10 (1.3)	7 (1.1)	4 (0.7)	13 (1.7)	11 (0.7)
California	14 (1.9)	18 (1.8)	6 (1.2)	6 (1.1)	11 (1.5)	11 (0.7)
Colorado	21 (1.5)	23 (1.7)	8 (1.2)	10 (1.0)	18 (1.6)	16 (0.7)
Connecticut	27 (1.9)	30 (2.4)	12 (1.6)	14 (1.7)	23 (1.6)	22 (1.0)
Delaware	21 (1.5)	20 (1.6)	7 (1.1)	11 (1.3)	17 (1.7)	15 (0.8)
Dist. Columbia	12 (1.6)	9 (1.0)	3 (0.8)	4 (1.1)	8 (1.6)	7 (0.7)
Florida	17 (1.4)	16 (1.6)	6 (1.2)	9 (1.2)	14 (1.8)	12 (0.8)
Georgia	25 (1.7)	16 (1.5)	8 (1.2)	8 (1.0)	19 (1.4)	15 (0.7)
Hawaii	17 (1.8)	20 (1.9)	7 (1.1)	7 (1.1)	12 (1.4)	12 (0.9)
Idaho	23 (2.0)	25 (1.9)	7 (1.0)	9 (1.3)	16 (1.3)	16 (0.8)
Indiana	24 (2.3)	21 (1.7)	6 (1.1)	8 (1.3)	18 (1.8)	15 (0.9)
Iowa	29 (1.6)	28 (2.2)	12 (1.3)	16 (1.9)	26 (1.9)	22 (1.0)
Kentucky	21 (2.0)	18 (1.7)	6 (1.2)	8 (1.5)	17 (1.5)	14 (0.9)
Louisiana	14 (1.5)	9 (1.5)	5 (0.8)	5 (1.0)	10 (1.6)	8 (0.7)
Maine	30 (2.3)	31 (2.9)	13 (1.6)	12 (1.7)	26 (2.4)	22 (1.0)
Maryland	21 (1.6)	23 (1.8)	9 (1.2)	13 (1.7)	20 (1.8)	17 (0.9)
Massachusetts	22 (2.3)	27 (2.5)	10 (1.6)	14 (1.3)	20 (1.8)	19 (1.1)
Michigan	21 (1.8)	22 (2.3)	10 (1.2)	7 (1.3)	19 (2.0)	16 (1.0)
Minnesota	27 (2.0)	26 (2.3)	11 (1.4)	9 (1.4)	24 (1.9)	20 (0.9)
Mississippi	11 (1.3)	10 (1.3)	4 (0.8)	7 (1.0)	8 (1.1)	8 (0.7)
Missouri	26 (2.1)	23 (2.0)	9 (1.3)	9 (1.4)	20 (1.8)	17 (0.9)
Nebraska	26 (2.5)	23 (1.9)	9 (1.1)	12 (1.7)	20 (1.8)	18 (0.9)
New Hampshire	28 (2.2)	29 (2.0)	9 (1.5)	14 (1.7)	26 (2.2)	21 (1.0)
New Jersey	22 (1.7)	27 (2.0)	10 (1.5)	9 (1.1)	23 (2.3)	18 (0.8)
New Mexico	17 (1.4)	18 (2.7)	4 (1.0)	9 (2.2)	14 (2.1)	12 (0.9)
New York	16 (1.7)	20 (1.8)	9 (1.5)	9 (1.5)	18 (1.9)	15 (1.0)
North Carolina	19 (1.5)	16 (1.7)	6 (0.9)	9 (1.3)	15 (1.5)	13 (0.7)
North Dakota	30 (2.0)	24 (1.7)	10 (1.1)	11 (1.5)	22 (2.0)	20 (0.8)
Ohio	24 (1.8)	23 (1.5)	9 (1.4)	11 (1.4)	19 (1.6)	17 (0.8)
Oklahoma	23 (1.8)	21 (1.8)	7 (1.0)	9 (1.1)	20 (1.7)	16 (0.7)
Pennsylvania	24 (1.7)	23 (1.6)	11 (1.7)	15 (1.4)	22 (1.7)	19 (0.9)
Rhode Island	22 (2.1)	18 (1.7)	7 (1.4)	12 (1.5)	18 (1.7)	16 (1.0)
South Carolina	17 (1.6)	13 (1.4)	6 (0.8)	8 (1.1)	13 (1.4)	11 (0.7)
Tennessee	23 (2.1)	16 (2.2)	6 (1.1)	8 (1.0)	13 (1.4)	13 (0.8)
Texas	17 (1.9)	21 (2.0)	11 (1.5)	9 (1.2)	14 (1.3)	15 (0.8)
Utah	23 (1.8)	21 (1.7)	6 (1.0)	10 (1.2)	19 (1.7)	16 (0.8)
Virginia	24 (1.6)	23 (1.9)	8 (1.6)	12 (1.4)	19 (1.7)	17 (1.0)
West Virginia	19 (1.7)	17 (1.7)	5 (1.0)	8 (1.1)	14 (1.5)	13 (0.7)
Wisconsin	25 (1.7)	29 (2.2)	11 (1.3)	13 (1.3)	24 (1.6)	20 (0.9)
Wyoming	25 (1.9)	24 (1.9)	11 (1.2)	10 (1.4)	21 (1.5)	18 (0.8)
TERRITORY						
Guam	7 (1.4)	8 (1.5)	3 (0.7)	5 (1.0)	10 (1.2)	7 (0.5)

*Secure question, unreleased. The standard errors of the estimated percentages appear in parentheses. It can be said with 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details).

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment.

BEST COPY AVAILABLE

TABLES 3.3 and 3.4 summarize performance at grade 8 for the nation and the states on the six extended-response questions included in the 1992 assessment. (Three of these tasks are found in Chapter Two.) National performance ranged from 4 to 13 percent of the students providing satisfactory or better responses, with an average of 8 percent. Across the states, average performance varied from 0 to 13 percent satisfactory or better.

TABLE 3.3 National Percentages of Satisfactory or Better Responses to Extended-Response Questions

Grade 8

	Treena's Budget	Radio Stations	Marcy's Dot Pattern	Probability (Leroy's Coins)*	Geometric Shapes (Hallway)*	Number Patterns (Tiles)*	Average Percent Satisfactory or Better
Nation	4 (0.5)	5 (0.6)	6 (0.7)	13 (1.1)	7 (0.7)	13 (1.1)	8 (0.5)
Northeast	4 (1.0)	6 (1.1)	8 (1.6)	17 (3.9)	8 (1.9)	16 (1.5)	10 (1.4)
Southeast	3 (0.8)	3 (0.8)	4 (1.3)	10 (1.2)	6 (1.1)	10 (1.8)	6 (0.6)
Central	6 (1.1)	6 (1.3)	7 (1.4)	12 (2.0)	8 (1.5)	14 (1.9)	9 (0.8)
West	3 (0.9)	6 (1.4)	4 (1.1)	14 (1.6)	7 (1.2)	14 (2.6)	8 (0.9)
White	5 (0.6)	7 (0.9)	8 (0.9)	16 (1.5)	9 (1.0)	16 (1.3)	10 (0.6)
Black	0 (0.5)	1 (0.6)	1 (0.5)	1 (0.5)	1 (0.9)	6 (1.7)	2 (0.3)
Hispanic	1 (0.6)	1 (0.6)	1 (0.5)	6 (1.9)	4 (1.3)	6 (1.6)	3 (0.5)
Male	2 (0.5)	4 (0.8)	5 (0.9)	12 (1.3)	7 (1.1)	10 (1.2)	7 (0.5)
Female	6 (0.8)	6 (1.0)	6 (0.9)	14 (1.7)	8 (1.0)	17 (1.8)	10 (0.7)
Advantaged Urban	7 (2.7)	10 (1.5)	13 (2.6)	28 (5.5)	16 (2.8)	20 (4.5)	16 (2.1)
Disadvantaged Urban	2 (1.1)	2 (1.5)	1 (0.7)	3 (1.2)	3 (1.2)	6 (2.4)	3 (0.8)
Extreme Rural	4 (1.8)	2 (1.2)	5 (2.3)	8 (2.1)	5 (1.8)	12 (2.6)	6 (1.0)
Other	4 (0.6)	5 (0.8)	5 (0.7)	12 (1.0)	7 (0.8)	14 (1.2)	8 (0.5)
Public	4 (0.5)	5 (0.7)	5 (0.6)	13 (1.2)	7 (0.8)	13 (1.1)	8 (0.5)
Catholic and Other Private	4 (1.0)	7 (1.4)	13 (2.0)	17 (2.3)	12 (1.9)	13 (1.9)	11 (1.0)

*Secure question, unreleased.

The standard errors of the estimated percentages appear in parentheses. It can be said with about 95 percent certainty for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details).

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment

TABLE 3.4 | **Percentage of Satisfactory or Better Responses to Extended-Response Questions, Grade 8**

PUBLIC SCHOOLS	Grade 8 - 1992						
	Treena's Budget	Radio Stations	Marcy's Dot Pattern	Probability* (Coins)	Geometric Shapes* (Hallway)	Number Patterns* (Tiles)	Average Percent Satisfactory or Better
NATION	4 (0.5)	5 (0.7)	5 (0.6)	13 (1.2)	7 (0.8)	13 (1.1)	8 (0.5)
Northeast	5 (1.3)	5 (1.6)	7 (1.8)	18 (4.3)	7 (2.3)	16 (2.1)	10 (1.7)
Southeast	3 (0.9)	3 (0.7)	3 (1.1)	10 (1.5)	5 (1.2)	10 (1.7)	6 (0.7)
Central	6 (1.2)	5 (1.4)	7 (1.4)	11 (2.2)	8 (1.7)	13 (2.0)	8 (0.9)
West	3 (1.0)	7 (1.5)	4 (1.1)	14 (1.7)	6 (1.2)	14 (2.7)	8 (0.9)
STATES							
Alabama	2 (0.7)	3 (0.8)	3 (0.7)	5 (1.0)	4 (0.8)	9 (1.5)	4 (0.5)
Arizona	4 (0.9)	3 (0.9)	6 (1.1)	11 (1.2)	7 (1.2)	11 (1.5)	7 (0.7)
Arkansas	2 (0.6)	3 (0.7)	4 (0.9)	8 (1.1)	4 (0.7)	9 (1.1)	5 (0.4)
California	3 (0.8)	4 (1.1)	6 (1.1)	12 (1.8)	5 (0.9)	12 (1.4)	7 (0.8)
Colorado	7 (0.9)	7 (1.1)	6 (1.1)	18 (2.1)	10 (1.2)	16 (1.7)	11 (0.6)
Connecticut	7 (1.0)	7 (1.2)	10 (1.3)	17 (1.7)	11 (1.2)	14 (1.1)	11 (0.6)
Delaware	4 (1.4)	4 (1.0)	5 (1.1)	9 (1.2)	5 (1.2)	10 (1.2)	6 (0.4)
Dist. Columbia	1 (0.7)	2 (0.7)	5 (0.9)	2 (0.4)	2 (0.9)	6 (1.0)	3 (0.4)
Florida	4 (0.9)	4 (0.9)	5 (0.9)	10 (1.2)	4 (0.8)	9 (1.4)	6 (0.4)
Georgia	2 (0.6)	5 (0.9)	5 (0.9)	10 (1.1)	4 (1.0)	9 (1.7)	6 (0.5)
Hawaii	1 (0.5)	3 (0.8)	4 (0.9)	9 (1.3)	4 (0.7)	8 (1.3)	5 (0.4)
Idaho	4 (0.7)	7 (1.2)	6 (1.3)	9 (1.1)	10 (1.3)	14 (1.5)	8 (0.6)
Indiana	3 (0.8)	5 (1.0)	6 (1.0)	15 (1.8)	9 (1.1)	15 (1.7)	9 (0.7)
Iowa	8 (1.1)	10 (1.2)	8 (1.1)	21 (2.0)	16 (1.5)	16 (1.7)	13 (0.8)
Kentucky	4 (0.9)	3 (0.6)	5 (1.1)	12 (1.5)	7 (1.3)	11 (1.4)	7 (0.5)
Louisiana	1 (0.5)	1 (0.5)	3 (0.8)	4 (1.0)	2 (0.7)	7 (1.2)	3 (0.5)
Maine	5 (1.1)	8 (1.1)	10 (1.5)	18 (1.4)	10 (1.5)	16 (1.6)	11 (0.7)
Maryland	4 (0.9)	5 (1.1)	9 (1.4)	15 (1.2)	8 (1.1)	18 (1.7)	10 (0.7)
Massachusetts	5 (1.0)	7 (1.0)	9 (1.4)	17 (1.8)	11 (1.3)	18 (1.9)	11 (0.7)
Michigan	4 (0.9)	7 (1.1)	7 (1.1)	14 (1.8)	8 (1.2)	14 (1.4)	9 (0.7)
Minnesota	6 (1.4)	11 (1.6)	9 (1.2)	21 (1.6)	15 (1.5)	12 (1.6)	13 (0.7)
Mississippi	1 (0.5)	2 (0.6)	3 (0.9)	7 (1.1)	2 (0.6)	7 (1.2)	4 (0.5)
Missouri	4 (0.7)	8 (1.2)	8 (1.2)	13 (1.4)	8 (1.3)	14 (1.6)	9 (0.7)
Nebraska	5 (0.9)	7 (1.1)	9 (1.4)	19 (2.0)	10 (1.6)	14 (1.5)	11 (0.7)
New Hampshire	5 (0.8)	8 (1.3)	9 (1.2)	19 (1.9)	11 (1.6)	15 (1.4)	11 (0.6)
New Jersey	5 (1.3)	6 (1.0)	10 (1.6)	15 (1.3)	11 (1.5)	16 (1.6)	10 (0.8)
New Mexico	1 (0.5)	3 (0.8)	4 (0.8)	8 (1.1)	5 (1.0)	9 (1.2)	5 (0.4)
New York	4 (1.0)	7 (1.2)	5 (1.0)	16 (1.9)	9 (1.3)	13 (1.9)	9 (0.7)
North Carolina	3 (0.8)	3 (0.7)	6 (0.9)	10 (1.3)	6 (1.0)	9 (1.5)	6 (0.5)
North Dakota	6 (1.2)	8 (1.5)	7 (1.1)	19 (2.2)	14 (1.6)	18 (1.8)	12 (0.8)
Ohio	5 (1.0)	5 (1.0)	6 (1.1)	14 (1.7)	8 (1.2)	17 (2.4)	9 (0.7)
Oklahoma	3 (0.8)	5 (1.1)	6 (1.3)	12 (1.4)	7 (1.6)	11 (1.5)	7 (0.6)
Pennsylvania	3 (0.8)	8 (1.5)	7 (1.3)	15 (1.4)	11 (1.8)	16 (1.8)	10 (0.9)
Rhode Island	6 (1.1)	5 (1.0)	7 (1.2)	10 (1.6)	5 (0.9)	14 (2.0)	8 (0.6)
South Carolina	2 (0.5)	3 (0.6)	5 (1.0)	12 (1.4)	6 (1.2)	10 (1.5)	7 (0.5)
Tennessee	3 (0.7)	2 (0.6)	3 (0.8)	8 (1.2)	5 (1.0)	10 (1.5)	5 (0.5)
Texas	4 (1.0)	5 (1.3)	6 (1.0)	11 (1.6)	8 (1.6)	10 (1.5)	8 (0.7)
Utah	4 (0.8)	8 (1.3)	6 (1.0)	14 (1.4)	9 (1.2)	13 (1.6)	9 (0.5)
Virginia	4 (0.9)	5 (1.0)	7 (1.1)	15 (1.7)	9 (1.5)	12 (1.3)	9 (0.6)
West Virginia	2 (0.7)	3 (0.6)	4 (0.8)	10 (1.4)	3 (0.7)	9 (1.1)	5 (0.4)
Wisconsin	6 (1.0)	8 (1.2)	11 (2.0)	18 (1.9)	10 (1.3)	15 (1.6)	11 (0.9)
Wyoming	5 (0.9)	8 (1.1)	7 (1.1)	15 (1.7)	10 (1.2)	8 (1.3)	9 (0.5)
TERRITORIES							
Guam	1 (0.4)	3 (0.8)	2 (0.8)	4 (1.5)	2 (0.8)	4 (1.1)	3 (0.5)
Virgin Islands	0 (0.0)	0 (0.3)	0 (0.0)	0 (0.0)	0 (0.3)	0 (0.0)	0 (0.1)

*Secure question, unreleased. The standard errors of the estimated percentages appear in parentheses. It can be said with 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details).

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment.

As shown in TABLE 3.5 summarizing twelfth graders' performance for the nation, the percentages of satisfactory or better responses ranged from 1 to 28 percent. On average, 9 percent of the high-school seniors provided responses judged as satisfactory or better.

TABLE 3.5 National Percentages of Satisfactory or Better Responses to Extended-Response Questions

Grade 12

	Effective Tax Rates	Patterns of Squares (ending in 5)	Graphing Path of Object	Bicycle Trip Graph*	Center of Disk*	Extend Pattern of Tiles*	Average Percent Satisfactory or Better
Nation	3 (0.7)	2 (0.4)	1 (0.3)	28 (1.5)	12 (1.0)	5 (0.6)	9 (0.4)
Northeast	5 (1.8)	4 (1.4)	1 (0.6)	31 (2.4)	11 (2.6)	4 (1.0)	10 (0.8)
Southeast	1 (0.4)	1 (0.6)	1 (0.4)	20 (2.3)	12 (1.6)	4 (0.6)	7 (0.6)
Central	4 (1.7)	2 (0.7)	2 (0.7)	31 (4.0)	14 (1.6)	7 (1.1)	10 (0.9)
West	3 (1.2)	1 (0.6)	1 (0.4)	28 (2.7)	10 (2.2)	6 (1.6)	8 (0.7)
White	4 (0.9)	3 (0.6)	2 (0.4)	32 (1.9)	14 (1.2)	6 (0.7)	10 (0.5)
Black	2 (1.7)	0 (0.0)	0 (0.0)	14 (2.6)	5 (2.3)	1 (0.8)	4 (0.7)
Hispanic	4 (0.4)	0 (0.0)	0 (0.0)	14 (2.7)	6 (2.2)	2 (1.6)	4 (0.6)
Male	3 (0.8)	2 (0.7)	2 (0.5)	26 (1.9)	11 (1.3)	6 (1.1)	8 (0.5)
Female	4 (1.1)	2 (0.5)	1 (0.3)	30 (2.4)	13 (1.6)	5 (0.8)	9 (0.5)
Advantaged Urban	9 (2.1)	6 (2.1)	2 (0.8)	38 (3.3)	14 (2.1)	8 (2.0)	13 (1.1)
Disadvantaged White	1 (0.8)	0 (0.0)	0 (0.0)	18 (2.7)	7 (2.2)	2 (0.7)	5 (0.7)
Extreme Rural	1 (0.4)	0 (0.0)	0 (0.4)	19 (3.9)	16 (2.4)	5 (1.3)	7 (0.8)
Other	3 (0.9)	2 (0.5)	2 (0.4)	29 (1.8)	12 (1.4)	6 (0.8)	9 (0.5)
Public	2 (0.6)	2 (0.4)	1 (0.3)	25 (1.7)	12 (1.3)	5 (0.7)	8 (0.4)
Catholic and Other Private	9 (2.1)	4 (1.1)	3 (1.0)	42 (2.5)	10 (2.1)	9 (2.1)	13 (1.0)

*Secure question, unreleased

The standard errors of the estimated percentages appear in parentheses. It can be said with about 95 percent certainty for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details). When the proportion of students is either 0 percent of 100 percent, the standard error is inestimable. However, percentages 99.5 percent and greater were rounded to 100 percent and percentages 0.5 or less were rounded to 0 percent.

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment

Across the three grades assessed, it is clear that students had great difficulty with these tasks. Because the generally low levels of performance preclude much variation, differences in performance among subgroups did not tend to be large. However, White students did outperform their Black and Hispanic counterparts at all three grades assessed. Also, students attending

schools in advantaged urban communities had higher percentages of success than students attending schools in disadvantaged urban communities.

TABLES 3.6 and 3.7 present corresponding summary analyses for the nation and states on the regular constructed-response questions (see Chapter One for examples of questions). Average performance on this type of short-response question was considerably higher than on the extended tasks. The average percentage correct by grade level was 42 percent for grade 4, 53 percent for grade 8, and 40 percent for grade 12. However, there was a range in performance across subgroups. For example, at grade 4, 47 percent of the White students, on average, provided correct responses to these types of questions, compared to 24 percent of the Black students, and 31 percent of Hispanic students. The differences between the two minority groups, as well as their differences with the majority group, were statistically significant.

Across the participating states and territories, the average percentage correct ranged from 27 to 51 percent at grade 4 and from 30 to 63 percent at grade 8.

TABLE 3.6 Average Percentage Correct for Regular Constructed-Response Questions

Grade 4

	Overall	Numbers & Operations	Measurement	Geometry	Data Analysis, Statistics, and Probability	Algebra and Functions
Nation	42 (0.5)	46 (0.5)	38 (0.7)	36 (0.7)	45 (0.6)	43 (0.8)
Northeast	45 (1.4)	50 (1.1)	41 (2.0)	38 (2.1)	49 (1.6)	48 (2.3)
Southeast	36 (1.2)	40 (1.3)	33 (1.5)	31 (1.1)	39 (1.7)	35 (1.9)
Central	45 (1.0)	48 (0.8)	42 (1.8)	38 (1.2)	48 (1.0)	46 (1.5)
West	42 (1.0)	46 (0.9)	36 (1.2)	37 (1.3)	43 (1.4)	42 (1.9)
White	47 (0.6)	50 (0.6)	43 (0.9)	41 (0.9)	51 (0.8)	49 (1.1)
Black	24 (0.8)	32 (1.0)	20 (1.2)	18 (1.0)	24 (1.3)	22 (1.3)
Hispanic	31 (0.7)	35 (0.7)	28 (1.3)	26 (1.1)	32 (1.2)	27 (1.4)
Male	43 (0.5)	46 (0.5)	38 (0.9)	38 (0.7)	45 (0.7)	43 (1.1)
Female	41 (0.7)	46 (0.7)	38 (1.0)	35 (0.8)	44 (0.8)	42 (1.0)
Advantaged Urban	54 (1.3)	56 (1.3)	51 (2.3)	45 (1.7)	59 (1.8)	56 (2.1)
Disadvantaged Urban	26 (1.4)	34 (1.6)	19 (1.8)	20 (1.4)	26 (1.8)	22 (2.4)
Extreme Rural	40 (2.6)	45 (2.4)	38 (3.6)	34 (2.9)	42 (2.9)	38 (3.5)
Other	42 (0.6)	46 (0.6)	38 (0.8)	37 (0.8)	45 (0.8)	44 (0.9)
Public	41 (0.6)	45 (0.6)	37 (0.8)	35 (0.7)	44 (0.7)	42 (0.9)
Catholic and Other Private	47 (0.9)	50 (0.9)	44 (1.1)	41 (1.4)	50 (1.3)	48 (1.8)

The standard errors of the estimated percentages appear in parentheses. It can be said with about 95 percent certainty for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details).

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment

TABLE 3.6 Average Percentage Correct for Regular Constructed-Response Questions (Continued)

Grade 8

	Overall	Numbers and Operations	Measurement	Geometry	Data Analysis, Statistics, and Probability	Algebra and Functions
Nation	53 (0.5)	60 (0.6)	44 (0.6)	58 (0.6)	46 (0.5)	55 (0.9)
Northeast	54 (1.5)	60 (1.6)	45 (1.5)	60 (1.5)	48 (1.6)	54 (3.0)
Southeast	49 (0.8)	56 (1.0)	40 (1.1)	53 (1.1)	42 (0.8)	51 (1.2)
Central	57 (1.1)	64 (1.4)	47 (1.1)	63 (1.2)	50 (1.1)	61 (1.5)
West	53 (1.0)	59 (1.0)	44 (1.4)	58 (1.2)	46 (0.8)	54 (1.6)
White	59 (0.6)	65 (0.6)	49 (0.7)	64 (0.7)	51 (0.5)	61 (1.0)
Black	36 (0.9)	44 (1.3)	27 (1.1)	39 (1.6)	31 (1.1)	37 (1.5)
Hispanic	42 (0.7)	48 (1.2)	35 (1.1)	46 (1.2)	34 (1.2)	39 (1.3)
Male	53 (0.7)	59 (0.9)	45 (0.8)	58 (0.8)	46 (0.6)	54 (1.0)
Female	54 (0.5)	61 (0.6)	43 (0.5)	59 (0.6)	46 (0.7)	56 (1.0)
Advantaged Urban	64 (1.8)	70 (1.9)	54 (1.6)	71 (1.6)	59 (2.0)	66 (3.4)
Disadvantaged Urban	37 (1.5)	45 (1.8)	30 (1.4)	40 (2.5)	31 (1.4)	37 (2.5)
Extreme Rural	53 (2.8)	62 (2.5)	43 (2.6)	56 (4.1)	46 (2.5)	57 (3.0)
Other	54 (0.6)	60 (0.7)	44 (0.8)	59 (0.8)	46 (0.5)	55 (0.9)
Public	52 (0.5)	59 (0.6)	43 (0.6)	58 (0.6)	45 (0.5)	54 (1.0)
Catholic and Other Private	60 (1.3)	66 (1.3)	49 (1.2)	65 (1.6)	54 (1.4)	60 (1.7)

Grade 12

	Overall	Numbers and Operations	Measurement	Geometry	Data Analysis, Statistics, and Probability	Algebra and Functions
Nation	40 (0.5)	49 (0.6)	26 (0.7)	40 (0.6)	47 (0.6)	35 (0.7)
Northeast	42 (0.9)	51 (1.0)	27 (1.0)	41 (1.0)	48 (1.3)	37 (1.3)
Southeast	36 (0.8)	47 (1.2)	20 (0.7)	34 (0.9)	43 (1.0)	30 (1.0)
Central	43 (1.2)	52 (1.4)	27 (1.3)	43 (1.0)	49 (1.3)	39 (1.8)
West	40 (1.0)	48 (1.1)	27 (1.8)	40 (1.3)	47 (0.8)	34 (1.4)
White	44 (0.6)	52 (0.7)	29 (0.8)	43 (0.7)	51 (0.7)	39 (0.8)
Black	26 (0.9)	38 (1.3)	11 (0.6)	23 (1.2)	34 (1.3)	20 (1.1)
Hispanic	32 (0.9)	41 (1.6)	19 (1.2)	33 (1.4)	38 (1.5)	24 (1.9)
Male	41 (0.7)	49 (0.8)	27 (0.9)	42 (0.8)	48 (0.8)	36 (0.9)
Female	40 (0.6)	50 (0.8)	24 (0.7)	38 (0.8)	47 (0.6)	34 (0.9)
Advantaged Urban	49 (1.8)	56 (2.0)	36 (2.2)	48 (2.1)	55 (2.0)	45 (2.3)
Disadvantaged Urban	30 (1.1)	38 (1.2)	16 (1.4)	29 (1.4)	37 (1.5)	22 (1.4)
Extreme Rural	37 (1.5)	48 (1.3)	23 (2.3)	36 (1.6)	44 (1.6)	30 (1.7)
Other	41 (0.6)	50 (0.7)	25 (0.7)	40 (0.7)	48 (0.7)	36 (0.8)
Public	39 (0.6)	48 (0.7)	24 (0.8)	39 (0.7)	46 (0.7)	34 (0.9)
Catholic and Other Private	48 (1.5)	56 (1.4)	33 (1.7)	46 (1.7)	54 (1.8)	45 (1.8)

TABLE 3.7 | **Average Percentage Correct for Regular Constructed-Response Questions**

PUBLIC SCHOOLS	Grade 4 - 1992					
	Overall	Numbers and Operations	Measurement	Geometry	Data Analysis, Statistics, and Probability	Algebra and Functions
NATION	41 (0.6)	45 (0.6)	37 (0.8)	35 (0.7)	44 (0.7)	42 (0.9)
Northeast	45 (1.5)	49 (1.2)	41 (2.1)	37 (2.3)	48 (1.8)	48 (2.5)
Southeast	36 (1.4)	39 (1.5)	32 (1.5)	30 (1.3)	38 (2.1)	34 (2.1)
Central	44 (1.2)	48 (1.1)	41 (2.0)	37 (1.2)	47 (1.2)	45 (1.8)
West	41 (1.1)	45 (1.0)	35 (1.2)	36 (1.3)	43 (1.5)	41 (2.0)
STATES						
Alabama	34 (1.0)	38 (1.0)	31 (1.1)	28 (0.9)	38 (1.2)	32 (1.5)
Arizona	40 (0.6)	43 (0.7)	35 (0.9)	36 (0.7)	42 (1.1)	41 (1.0)
Arkansas	36 (0.6)	40 (0.6)	31 (0.9)	29 (0.8)	40 (1.0)	33 (1.0)
California	37 (0.9)	41 (0.9)	32 (1.1)	33 (1.2)	37 (1.4)	39 (1.2)
Colorado	44 (0.7)	46 (0.7)	40 (1.0)	40 (0.9)	47 (1.0)	45 (1.1)
Connecticut	48 (0.8)	50 (0.8)	44 (0.9)	43 (1.0)	52 (1.2)	48 (1.4)
Delaware	40 (0.6)	45 (0.8)	36 (0.7)	34 (0.7)	44 (1.0)	40 (1.0)
Dist. Columbia	27 (0.3)	35 (0.5)	20 (0.7)	23 (0.6)	25 (0.7)	25 (0.8)
Florida	39 (1.1)	42 (1.0)	35 (1.5)	32 (1.1)	43 (1.4)	40 (1.6)
Georgia	39 (0.9)	43 (0.9)	34 (1.1)	32 (0.9)	44 (1.2)	39 (1.6)
Hawaii	40 (0.7)	44 (0.8)	33 (0.9)	36 (0.9)	41 (0.9)	38 (1.0)
Idaho	43 (0.6)	45 (0.7)	41 (0.9)	40 (0.7)	45 (1.0)	45 (1.3)
Indiana	42 (0.7)	46 (0.8)	39 (0.9)	36 (1.0)	47 (1.1)	43 (1.2)
Iowa	49 (0.7)	52 (0.8)	45 (0.8)	42 (0.8)	55 (1.0)	51 (1.3)
Kentucky	38 (0.6)	43 (0.7)	33 (0.8)	31 (0.8)	43 (1.1)	38 (1.2)
Louisiana	33 (0.8)	38 (0.8)	29 (1.1)	27 (0.8)	34 (1.2)	30 (1.3)
Maine	51 (0.7)	51 (0.8)	45 (1.2)	49 (0.8)	55 (1.3)	53 (1.1)
Maryland	42 (0.7)	46 (0.8)	36 (0.8)	36 (0.8)	47 (1.2)	42 (1.0)
Massachusetts	47 (0.9)	50 (1.0)	42 (1.0)	42 (0.9)	51 (1.4)	46 (1.4)
Michigan	42 (1.1)	45 (1.0)	40 (1.3)	38 (1.2)	45 (1.4)	42 (1.6)
Minnesota	48 (0.7)	51 (0.7)	45 (1.0)	43 (0.9)	53 (1.2)	49 (1.2)
Mississippi	30 (0.6)	37 (0.7)	27 (0.8)	22 (0.7)	30 (0.9)	26 (1.0)
Missouri	44 (0.8)	47 (0.7)	40 (1.2)	37 (0.9)	49 (1.2)	44 (1.1)
Nebraska	46 (0.9)	49 (0.8)	42 (1.3)	42 (0.9)	50 (1.3)	46 (1.4)
New Hampshire	49 (0.9)	51 (0.8)	45 (1.1)	45 (1.3)	54 (1.2)	53 (1.3)
New Jersey	47 (1.0)	50 (1.0)	43 (1.2)	39 (1.1)	50 (1.3)	49 (1.4)
New Mexico	39 (1.0)	42 (0.9)	33 (1.2)	35 (1.5)	42 (1.2)	36 (1.6)
New York	42 (0.8)	46 (0.7)	36 (1.2)	33 (1.0)	47 (1.2)	41 (1.2)
North Carolina	38 (0.7)	42 (0.7)	32 (0.8)	31 (0.9)	41 (0.9)	38 (1.1)
North Dakota	48 (0.6)	50 (0.6)	45 (1.0)	41 (0.9)	55 (0.9)	48 (1.1)
Ohio	42 (0.8)	45 (0.8)	38 (1.1)	35 (1.0)	46 (1.1)	42 (1.1)
Oklahoma	42 (0.8)	46 (0.8)	37 (0.9)	34 (1.1)	47 (1.2)	42 (1.1)
Pennsylvania	45 (0.9)	49 (0.8)	42 (1.2)	37 (1.0)	50 (1.3)	45 (1.2)
Rhode Island	39 (0.9)	43 (0.9)	35 (1.1)	33 (1.2)	41 (1.2)	38 (1.4)
South Carolina	37 (0.8)	41 (0.6)	33 (1.1)	31 (0.9)	39 (1.1)	35 (1.3)
Tennessee	36 (0.8)	41 (0.8)	31 (1.0)	28 (0.8)	39 (1.3)	36 (1.2)
Texas	40 (0.9)	44 (0.9)	37 (1.1)	34 (1.0)	45 (1.3)	40 (1.4)
Utah	44 (0.7)	47 (0.7)	41 (1.1)	40 (0.9)	46 (1.0)	46 (1.2)
Virginia	43 (0.9)	46 (0.8)	37 (1.2)	37 (1.2)	49 (1.2)	43 (1.3)
West Virginia	39 (0.6)	42 (0.7)	36 (0.9)	33 (0.8)	42 (1.0)	39 (1.1)
Wisconsin	48 (0.8)	50 (0.8)	43 (1.1)	41 (1.0)	55 (1.0)	50 (1.3)
Wyoming	46 (0.6)	49 (0.6)	41 (0.9)	41 (0.8)	50 (1.0)	47 (1.1)
TERRITORY						
Guam	27 (0.5)	33 (0.5)	20 (0.6)	26 (0.7)	26 (0.7)	25 (1.1)

The standard errors of the estimated percentages appear in parentheses. It can be said with 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details).

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment.

BEST COPY AVAILABLE

TABLE 3.7 | **Average Percentage Correct for Regular Constructed-Response Questions (continued)**

PUBLIC SCHOOLS	Grade 8 - 1992					
	Overall	Numbers and Operations	Measurement	Geometry	Data Analysis, Statistics, and Probability	Algebra and Functions
NATION	52 (0.5)	59 (0.6)	43 (0.6)	57 (0.6)	45 (0.5)	54 (1.0)
Northeast	53 (1.8)	59 (1.9)	44 (1.4)	59 (1.8)	47 (1.9)	52 (3.6)
Southeast	48 (0.7)	55 (0.9)	40 (1.2)	51 (1.0)	40 (0.6)	50 (1.4)
Central	57 (1.2)	64 (1.7)	46 (1.2)	62 (1.3)	49 (1.2)	60 (1.7)
West	52 (1.1)	58 (1.1)	44 (1.4)	58 (1.2)	46 (0.9)	54 (1.6)
STATES						
Alabama	46 (0.9)	53 (0.9)	38 (1.0)	50 (1.1)	40 (0.8)	46 (1.2)
Arizona	53 (0.7)	59 (0.8)	45 (0.8)	59 (0.9)	45 (0.7)	54 (0.9)
Arkansas	48 (0.7)	55 (0.9)	40 (0.6)	53 (1.1)	41 (0.8)	48 (1.0)
California	51 (0.9)	56 (1.1)	44 (0.9)	59 (1.1)	43 (1.0)	52 (1.2)
Colorado	58 (0.6)	63 (0.8)	49 (0.7)	64 (0.8)	50 (0.6)	60 (0.8)
Connecticut	58 (0.7)	64 (0.8)	50 (0.7)	64 (0.8)	50 (0.8)	58 (1.1)
Delaware	53 (0.6)	59 (0.9)	43 (0.7)	58 (0.9)	46 (0.8)	54 (1.0)
Dist. Columbia	37 (0.5)	44 (0.8)	29 (0.8)	39 (1.0)	31 (0.7)	37 (0.8)
Florida	51 (0.8)	57 (0.8)	42 (0.9)	55 (1.0)	44 (0.9)	53 (1.1)
Georgia	49 (0.7)	56 (0.8)	40 (0.8)	54 (1.0)	44 (0.9)	50 (1.1)
Hawaii	50 (0.5)	56 (0.7)	42 (0.7)	58 (0.8)	39 (0.7)	48 (1.0)
Idaho	59 (0.5)	65 (0.6)	50 (0.8)	67 (0.7)	50 (0.6)	62 (0.8)
Indiana	56 (0.6)	62 (0.8)	47 (0.8)	63 (0.8)	49 (0.7)	56 (0.8)
Iowa	63 (0.6)	69 (0.8)	54 (0.6)	68 (0.8)	55 (0.6)	65 (0.9)
Kentucky	51 (0.6)	58 (0.7)	42 (0.7)	56 (0.8)	45 (0.9)	52 (1.1)
Louisiana	44 (0.9)	51 (1.1)	36 (0.9)	48 (1.2)	38 (1.0)	46 (1.0)
Maine	62 (0.5)	66 (0.7)	53 (0.6)	69 (0.8)	54 (0.8)	63 (0.8)
Maryland	53 (0.7)	58 (0.8)	43 (0.7)	59 (1.1)	47 (0.8)	55 (1.1)
Massachusetts	58 (0.6)	65 (0.9)	48 (0.7)	64 (0.9)	50 (0.7)	61 (1.0)
Michigan	55 (0.8)	61 (0.8)	46 (0.9)	60 (1.1)	48 (0.8)	57 (1.0)
Minnesota	63 (0.5)	68 (0.6)	55 (0.8)	69 (0.7)	54 (0.7)	65 (1.0)
Mississippi	42 (0.7)	52 (0.9)	34 (0.8)	45 (0.9)	37 (0.9)	41 (1.1)
Missouri	57 (0.7)	62 (0.8)	48 (0.8)	64 (0.9)	50 (0.9)	59 (1.0)
Nebraska	60 (0.6)	65 (0.8)	51 (0.8)	67 (0.9)	51 (0.7)	62 (1.1)
New Hampshire	61 (0.5)	66 (0.7)	52 (0.8)	68 (0.7)	53 (0.7)	63 (0.8)
New Jersey	57 (0.8)	64 (0.9)	48 (0.9)	61 (1.0)	50 (0.9)	57 (1.2)
New Mexico	50 (0.6)	56 (0.7)	42 (0.7)	57 (0.8)	42 (0.7)	50 (1.1)
New York	54 (1.1)	61 (1.2)	46 (1.1)	60 (1.4)	47 (1.2)	55 (1.6)
North Carolina	50 (0.6)	55 (0.7)	41 (0.8)	55 (0.9)	44 (0.7)	50 (0.9)
North Dakota	63 (0.7)	69 (0.8)	54 (0.7)	70 (0.9)	55 (0.8)	67 (1.2)
Ohio	55 (0.8)	62 (0.9)	45 (0.9)	60 (1.1)	48 (0.8)	57 (1.3)
Oklahoma	55 (0.6)	62 (0.9)	45 (0.6)	60 (0.7)	48 (0.8)	57 (1.0)
Pennsylvania	57 (0.8)	63 (1.0)	49 (0.8)	63 (1.0)	49 (1.0)	60 (1.3)
Rhode Island	54 (0.4)	60 (0.6)	45 (0.5)	59 (0.7)	47 (0.7)	56 (1.5)
South Carolina	50 (0.6)	56 (0.8)	42 (0.7)	56 (0.8)	43 (0.6)	49 (0.9)
Tennessee	50 (0.8)	57 (1.0)	41 (0.9)	54 (1.0)	44 (0.8)	49 (1.1)
Texas	52 (0.7)	57 (0.8)	44 (0.8)	59 (0.9)	46 (0.8)	54 (1.1)
Utah	59 (0.5)	64 (0.7)	49 (0.6)	66 (0.7)	51 (0.6)	62 (0.8)
Virginia	55 (0.7)	62 (0.9)	45 (0.9)	60 (0.9)	47 (0.7)	56 (1.0)
West Virginia	50 (0.6)	56 (0.8)	42 (0.7)	55 (0.9)	43 (0.6)	51 (1.0)
Wisconsin	60 (0.8)	66 (1.0)	51 (0.7)	66 (0.8)	53 (1.0)	63 (1.4)
Wyoming	59 (0.5)	64 (0.6)	50 (0.5)	66 (0.6)	51 (0.6)	61 (1.0)
TERRITORIES						
Guam	38 (0.6)	42 (0.8)	32 (0.8)	46 (1.1)	29 (0.8)	37 (0.9)
Virgin Islands	30 (0.4)	37 (0.7)	24 (0.6)	34 (0.8)	24 (0.7)	27 (0.8)

The results of parallel analyses for the multiple-choice questions are found in TABLES 3.8 and 3.9. These results are similar to those for the regular constructed-response questions, although average performance for the nation was even higher, especially at grades 4 and 12.

TABLE 3.8 Average Percentage Correct for Multiple-Choice Questions

Grade 4

	Overall	Numbers and Operations	Measurement	Geometry	Data Analysis, Statistics, and Probability	Algebra and Functions
Nation	50 (0.4)	51 (0.5)	50 (0.5)	53 (0.4)	46 (0.5)	42 (0.5)
Northeast	52 (1.1)	54 (1.1)	52 (1.1)	55 (1.3)	47 (1.5)	46 (1.1)
Southeast	46 (0.8)	46 (0.9)	46 (0.9)	50 (1.0)	43 (1.1)	38 (1.3)
Central	51 (1.0)	52 (1.0)	53 (1.2)	55 (1.0)	48 (1.3)	43 (1.2)
West	49 (0.8)	50 (1.0)	50 (0.9)	53 (0.6)	46 (1.0)	42 (0.9)
White	53 (0.5)	54 (0.5)	55 (0.5)	56 (0.5)	50 (0.7)	44 (0.6)
Black	38 (0.6)	39 (0.8)	36 (0.8)	43 (0.9)	32 (1.0)	31 (1.1)
Hispanic	42 (0.7)	43 (0.9)	41 (0.9)	47 (1.0)	37 (1.4)	36 (1.0)
Male	50 (0.5)	51 (0.5)	53 (0.6)	54 (0.6)	47 (0.8)	42 (0.7)
Female	48 (0.5)	50 (0.6)	48 (0.6)	52 (0.6)	44 (0.8)	41 (0.8)
Advantaged Urban	59 (1.4)	61 (1.6)	60 (1.5)	55 (1.0)	56 (1.9)	51 (1.6)
Disadvantaged Urban	38 (1.2)	40 (1.3)	36 (1.3)	50 (1.0)	33 (0.9)	32 (1.6)
Extreme Rural	48 (1.7)	49 (1.8)	49 (2.1)	55 (1.3)	46 (1.8)	40 (1.5)
Other	50 (0.5)	50 (0.5)	51 (0.5)	53 (0.6)	46 (0.6)	42 (0.7)
Public	49 (0.5)	50 (0.5)	50 (0.5)	53 (0.5)	45 (0.6)	41 (0.6)
Catholic and Other Private	53 (0.8)	55 (0.9)	54 (0.9)	54 (0.9)	50 (1.0)	45 (1.1)

(Table 3.8 continued on the next page)

The standard errors of the estimated percentages appear in parentheses. It can be said with about 95 percent certainty for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details).

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment

Table 3.8 Average Percentage Correct for Multiple-Choice Questions (continued)

Grade 8

	Overall	Numbers and Operations	Measurement	Geometry	Data Analysis, Statistics, and Probability	Algebra and Functions
Nation	56 (0.4)	64 (0.5)	57 (0.5)	49 (0.5)	51 (0.6)	51 (0.5)
Northeast	57 (1.3)	65 (1.1)	57 (1.2)	50 (1.5)	52 (1.7)	52 (1.5)
Southeast	53 (0.7)	61 (0.8)	53 (0.9)	46 (0.6)	47 (1.1)	48 (0.9)
Central	59 (0.8)	68 (0.8)	60 (1.3)	51 (0.9)	53 (0.9)	53 (1.4)
West	56 (1.0)	64 (1.1)	58 (1.2)	49 (1.0)	51 (1.4)	51 (1.1)
White	60 (0.5)	69 (0.5)	62 (0.6)	52 (0.5)	55 (0.7)	55 (0.7)
Black	42 (0.6)	51 (0.7)	41 (0.8)	36 (0.8)	36 (1.0)	38 (0.8)
Hispanic	46 (0.7)	54 (0.9)	48 (0.8)	41 (0.7)	39 (0.9)	41 (1.0)
Male	56 (0.6)	64 (0.6)	58 (0.7)	49 (0.7)	51 (0.8)	50 (0.8)
Female	56 (0.5)	64 (0.5)	56 (0.6)	48 (0.6)	51 (0.6)	52 (0.7)
Advantaged Urban	65 (1.9)	73 (1.7)	64 (1.6)	58 (2.0)	60 (2.5)	63 (2.3)
Disadvantaged Urban	43 (1.2)	50 (1.4)	42 (1.5)	38 (1.1)	35 (1.6)	38 (1.4)
Extreme Rural	56 (2.3)	65 (2.2)	57 (2.3)	49 (2.1)	52 (3.0)	50 (2.3)
Other	56 (0.5)	64 (0.6)	58 (0.6)	49 (0.5)	51 (0.7)	51 (0.7)
Public	56 (0.5)	64 (0.5)	56 (0.5)	48 (0.5)	50 (0.7)	50 (0.6)
Catholic and Other Private	62 (1.1)	70 (1.1)	62 (1.1)	54 (1.1)	56 (1.0)	58 (1.6)

Grade 12

	Overall	Numbers and Operations	Measurement	Geometry	Data Analysis, Statistics, and Probability	Algebra and Functions
Nation	56 (0.4)	64 (0.4)	60 (0.5)	57 (0.6)	54 (0.4)	48 (0.5)
Northeast	58 (0.6)	65 (0.6)	61 (0.8)	59 (0.9)	55 (0.8)	51 (0.9)
Southeast	54 (0.6)	63 (0.7)	55 (0.7)	54 (1.1)	52 (0.8)	45 (0.8)
Central	58 (0.9)	65 (0.8)	63 (1.2)	59 (1.0)	56 (0.7)	50 (1.2)
West	56 (0.8)	64 (0.7)	60 (1.2)	57 (1.2)	55 (0.9)	47 (1.1)
White	59 (0.4)	67 (0.4)	63 (0.6)	60 (0.5)	57 (0.5)	50 (0.6)
Black	46 (0.9)	54 (0.8)	48 (1.2)	45 (1.2)	43 (1.1)	39 (1.1)
Hispanic	49 (1.0)	56 (1.0)	51 (0.9)	48 (1.9)	50 (1.3)	40 (1.3)
Male	58 (0.5)	65 (0.5)	62 (0.7)	59 (0.8)	56 (0.6)	49 (0.7)
Female	55 (0.5)	63 (0.5)	58 (0.7)	55 (0.7)	53 (0.5)	48 (0.6)
Advantaged Urban	65 (1.1)	72 (0.8)	69 (1.5)	68 (1.8)	61 (1.0)	59 (1.4)
Disadvantaged Urban	48 (1.2)	56 (1.1)	50 (1.6)	49 (1.6)	48 (1.3)	40 (1.2)
Extreme Rural	53 (1.2)	62 (0.9)	57 (1.9)	53 (1.2)	53 (1.4)	44 (1.5)
Other	57 (0.4)	64 (0.4)	60 (0.6)	57 (0.7)	54 (0.6)	48 (0.6)
Public	56 (0.5)	63 (0.4)	59 (0.6)	56 (0.6)	54 (0.5)	47 (0.6)
Catholic and Other Private	63 (1.2)	69 (1.0)	67 (1.5)	65 (1.6)	59 (1.2)	56 (1.3)

TABLE 3.9 | **Average Percentage Correct for Multiple-Choice Questions**

PUBLIC SCHOOLS	Grade 4 - 1992					
	Overall	Numbers and Operations	Measurement	Geometry	Data Analysis, Statistics, and Probability	Algebra and Functions
NATION	49 (0.5)	50 (0.5)	50 (0.5)	53 (0.5)	45 (0.6)	41 (0.6)
Northeast	52 (1.1)	54 (1.2)	51 (1.2)	54 (1.4)	46 (1.6)	45 (1.2)
Southeast	45 (0.9)	46 (1.0)	46 (1.0)	49 (1.1)	42 (1.2)	37 (1.4)
Central	51 (1.2)	52 (1.1)	52 (1.5)	55 (1.2)	47 (1.5)	42 (1.2)
West	49 (0.8)	50 (1.1)	50 (1.0)	53 (0.6)	45 (1.0)	41 (0.9)
STATES						
Alabama	45 (0.7)	47 (0.8)	45 (0.8)	48 (0.7)	41 (0.9)	38 (0.9)
Arizona	47 (0.5)	49 (0.6)	48 (0.5)	50 (0.7)	42 (0.7)	41 (0.7)
Arkansas	46 (0.4)	47 (0.5)	47 (0.5)	50 (0.7)	40 (0.8)	38 (0.5)
California	45 (0.7)	47 (0.7)	46 (0.8)	49 (0.7)	40 (1.0)	38 (0.9)
Colorado	50 (0.5)	51 (0.5)	51 (0.7)	56 (0.6)	45 (0.8)	41 (0.6)
Connecticut	53 (0.6)	55 (0.6)	53 (0.7)	55 (0.7)	48 (1.0)	48 (0.8)
Delaware	49 (0.4)	51 (0.4)	49 (0.6)	52 (0.6)	45 (0.8)	42 (0.8)
Dist. Columbia	39 (0.3)	39 (0.4)	37 (0.5)	45 (0.5)	35 (0.7)	34 (0.6)
Florida	47 (0.7)	48 (0.7)	49 (0.9)	50 (0.7)	43 (0.8)	40 (0.9)
Georgia	49 (0.7)	49 (0.8)	49 (0.7)	53 (0.7)	46 (1.0)	43 (0.8)
Hawaii	48 (0.6)	49 (0.6)	48 (0.7)	51 (0.7)	42 (1.0)	42 (0.7)
Idaho	50 (0.5)	51 (0.5)	52 (0.6)	54 (0.8)	44 (0.8)	41 (0.6)
Indiana	50 (0.5)	51 (0.6)	53 (0.6)	53 (0.7)	46 (0.8)	42 (0.6)
Iowa	54 (0.6)	56 (0.6)	56 (0.8)	56 (0.7)	49 (0.8)	46 (0.8)
Kentucky	48 (0.5)	49 (0.6)	49 (0.6)	50 (0.6)	43 (0.8)	41 (0.6)
Louisiana	43 (0.6)	45 (0.6)	43 (0.8)	46 (0.8)	39 (0.7)	36 (0.7)
Maine	56 (0.6)	56 (0.6)	58 (0.8)	58 (0.8)	53 (0.9)	48 (0.7)
Maryland	49 (0.6)	50 (0.6)	49 (0.8)	53 (0.7)	44 (0.9)	41 (0.7)
Massachusetts	53 (0.6)	54 (0.7)	54 (0.8)	55 (0.8)	49 (1.0)	46 (0.9)
Michigan	50 (0.8)	51 (0.8)	52 (0.9)	54 (0.8)	45 (0.8)	43 (0.9)
Minnesota	54 (0.4)	56 (0.5)	56 (0.6)	57 (0.7)	49 (0.7)	46 (0.6)
Mississippi	42 (0.5)	44 (0.6)	42 (0.6)	47 (0.5)	37 (0.9)	35 (0.7)
Missouri	51 (0.6)	52 (0.6)	52 (0.8)	55 (0.7)	46 (0.8)	43 (0.7)
Nebraska	53 (0.7)	54 (0.7)	54 (0.8)	56 (0.8)	47 (0.8)	44 (1.0)
New Hampshire	54 (0.7)	55 (0.7)	56 (0.8)	57 (0.8)	49 (0.9)	45 (1.0)
New Jersey	54 (0.7)	56 (0.7)	54 (1.0)	56 (0.7)	49 (1.0)	47 (0.7)
New Mexico	46 (0.7)	47 (0.8)	47 (0.9)	51 (0.8)	41 (0.9)	39 (1.1)
New York	50 (0.5)	51 (0.6)	49 (0.7)	53 (0.6)	50 (0.8)	42 (0.8)
North Carolina	47 (0.5)	48 (0.6)	47 (0.6)	51 (0.6)	43 (0.8)	40 (0.6)
North Dakota	54 (0.6)	55 (0.6)	56 (0.7)	56 (0.8)	49 (0.8)	46 (0.8)
Ohio	49 (0.6)	51 (0.6)	49 (0.8)	53 (0.7)	44 (0.9)	42 (0.8)
Oklahoma	50 (0.5)	51 (0.6)	51 (0.7)	52 (0.7)	45 (0.9)	43 (0.7)
Pennsylvania	52 (0.7)	54 (0.7)	54 (0.9)	54 (0.8)	47 (0.8)	45 (0.9)
Rhode Island	48 (0.7)	50 (0.7)	49 (0.8)	50 (0.7)	42 (0.9)	41 (0.8)
South Carolina	47 (0.6)	48 (0.6)	48 (0.7)	51 (0.7)	42 (0.8)	39 (0.8)
Tennessee	46 (0.6)	48 (0.6)	45 (0.7)	50 (0.7)	42 (1.0)	39 (0.6)
Texas	50 (0.6)	51 (0.6)	50 (0.7)	54 (0.8)	42 (0.9)	44 (0.8)
Utah	52 (0.5)	53 (0.6)	54 (0.6)	55 (0.7)	48 (0.8)	44 (0.8)
Virginia	50 (0.6)	52 (0.7)	51 (0.7)	54 (0.6)	47 (1.0)	42 (0.9)
West Virginia	48 (0.5)	48 (0.6)	50 (0.6)	51 (0.6)	42 (0.7)	39 (0.8)
Wisconsin	54 (0.6)	56 (0.7)	57 (0.7)	56 (0.7)	49 (0.8)	46 (0.7)
Wyoming	52 (0.5)	53 (0.5)	53 (0.6)	55 (0.6)	46 (0.7)	43 (0.7)
TERRITORY						
Guam	39 (0.3)	40 (0.4)	38 (0.6)	45 (0.7)	34 (0.8)	35 (0.7)

The standard errors of the estimated percentages appear in parentheses. It can be said with 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details).

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment.

TABLE 3.9 | **Average Percentage Correct for Multiple-Choice Questions (continued)**

PUBLIC SCHOOLS	Grade 8 - 1992					
	Overall	Numbers and Operations	Measurement	Geometry	Data Analysis, Statistics, and Probability	Algebra and Functions
NATION	56 (0.5)	64 (0.5)	57 (0.5)	48 (0.5)	50 (0.7)	50 (0.6)
Northeast	56 (1.4)	63 (1.2)	57 (1.4)	50 (1.5)	51 (2.0)	52 (1.5)
Southeast	52 (0.7)	60 (0.8)	52 (0.8)	44 (0.5)	46 (1.0)	46 (0.9)
Central	59 (1.0)	67 (1.0)	60 (1.2)	51 (1.0)	53 (1.1)	53 (1.7)
West	56 (1.1)	64 (1.2)	57 (1.2)	48 (1.0)	50 (1.5)	51 (1.2)
STATES						
Alabama	51 (0.7)	59 (0.7)	51 (0.7)	42 (0.7)	44 (0.9)	46 (0.8)
Arizona	56 (0.5)	64 (0.6)	57 (0.6)	48 (0.6)	50 (0.6)	50 (0.7)
Arkansas	52 (0.5)	61 (0.6)	52 (0.7)	45 (0.6)	46 (0.6)	46 (0.7)
California	54 (0.8)	62 (0.8)	55 (0.8)	48 (0.9)	47 (1.0)	49 (0.9)
Colorado	58 (0.5)	66 (0.5)	60 (0.7)	52 (0.6)	54 (0.7)	53 (0.6)
Connecticut	60 (0.5)	68 (0.6)	61 (0.6)	52 (0.6)	54 (0.6)	54 (0.7)
Delaware	54 (0.5)	63 (0.6)	55 (0.8)	47 (0.6)	49 (0.7)	49 (0.7)
Dist. Columbia	44 (0.4)	52 (0.4)	43 (0.7)	39 (0.7)	35 (0.7)	40 (0.4)
Florida	53 (0.7)	61 (0.8)	53 (0.7)	46 (0.7)	48 (0.9)	48 (0.8)
Georgia	54 (0.5)	62 (0.6)	53 (0.7)	46 (0.6)	47 (0.7)	48 (0.7)
Hawaii	52 (0.5)	60 (0.6)	53 (0.5)	47 (0.6)	44 (0.6)	48 (0.7)
Idaho	60 (0.4)	68 (0.5)	61 (0.6)	53 (0.5)	54 (0.5)	55 (0.6)
Indiana	58 (0.5)	65 (0.6)	59 (0.6)	51 (0.7)	53 (0.6)	52 (0.7)
Iowa	64 (0.6)	72 (0.6)	66 (0.6)	57 (0.7)	59 (0.6)	58 (0.7)
Kentucky	54 (0.5)	62 (0.6)	55 (0.6)	47 (0.5)	49 (0.7)	49 (0.7)
Louisiana	49 (0.7)	58 (0.8)	49 (0.8)	42 (0.8)	43 (0.9)	44 (0.8)
Maine	61 (0.5)	69 (0.6)	63 (0.8)	54 (0.5)	58 (0.6)	54 (0.7)
Maryland	56 (0.6)	64 (0.7)	56 (0.7)	49 (0.7)	52 (0.8)	51 (0.9)
Massachusetts	59 (0.5)	68 (0.6)	59 (0.5)	51 (0.7)	54 (0.7)	54 (0.7)
Michigan	57 (0.6)	65 (0.8)	57 (0.7)	50 (0.8)	51 (0.7)	52 (0.9)
Minnesota	64 (0.5)	71 (0.5)	64 (0.6)	57 (0.5)	59 (0.7)	59 (0.7)
Mississippi	48 (0.5)	58 (0.5)	48 (0.6)	39 (0.6)	42 (0.7)	43 (0.6)
Missouri	58 (0.6)	66 (0.6)	59 (0.6)	52 (0.7)	53 (0.8)	52 (0.7)
Nebraska	62 (0.6)	69 (0.6)	62 (0.6)	55 (0.8)	56 (0.7)	56 (0.8)
New Hampshire	61 (0.5)	69 (0.5)	62 (0.6)	54 (0.7)	56 (0.6)	54 (0.7)
New Jersey	59 (0.7)	68 (0.7)	58 (0.9)	52 (0.8)	52 (0.9)	55 (0.9)
New Mexico	53 (0.5)	61 (0.6)	54 (0.6)	46 (0.6)	47 (0.5)	47 (0.6)
New York	57 (0.9)	65 (0.9)	56 (0.8)	50 (1.1)	53 (1.1)	52 (1.0)
North Carolina	53 (0.6)	61 (0.6)	52 (0.7)	46 (0.6)	47 (0.7)	48 (0.7)
North Dakota	64 (0.5)	73 (0.5)	65 (0.7)	57 (0.7)	60 (0.7)	58 (0.7)
Ohio	57 (0.7)	65 (0.7)	58 (1.0)	49 (0.7)	52 (0.7)	51 (0.8)
Oklahoma	56 (0.5)	64 (0.6)	58 (0.7)	48 (0.7)	51 (0.7)	52 (0.6)
Pennsylvania	58 (0.7)	66 (0.8)	59 (0.7)	51 (0.8)	53 (0.8)	53 (0.7)
Rhode Island	56 (0.4)	64 (0.5)	56 (0.7)	48 (0.6)	50 (0.7)	50 (0.6)
South Carolina	54 (0.4)	62 (0.5)	54 (0.6)	48 (0.6)	47 (0.6)	49 (0.6)
Tennessee	53 (0.7)	62 (0.7)	53 (0.7)	45 (0.7)	48 (0.8)	47 (0.9)
Texas	56 (0.6)	63 (0.7)	55 (0.7)	50 (0.7)	50 (0.8)	52 (0.7)
Utah	60 (0.4)	67 (0.5)	61 (0.4)	52 (0.6)	56 (0.6)	54 (0.7)
Virginia	57 (0.6)	66 (0.6)	57 (0.7)	49 (0.7)	51 (0.7)	52 (0.7)
West Virginia	52 (0.5)	60 (0.6)	54 (0.6)	45 (0.6)	47 (0.6)	46 (0.7)
Wisconsin	62 (0.7)	70 (0.7)	63 (0.7)	55 (0.7)	57 (0.9)	56 (0.8)
Wyoming	60 (0.5)	68 (0.6)	61 (0.5)	54 (0.6)	54 (0.6)	53 (0.6)
TERRITORIES						
Guam	44 (0.5)	52 (0.6)	44 (0.7)	40 (0.7)	34 (0.7)	41 (0.6)
Virgin Islands	38 (0.3)	46 (0.5)	39 (0.6)	33 (0.5)	29 (0.6)	34 (0.6)

The Information Provided by Constructed-Response Tasks

Chapters One and Two of this report provide evidence of the rich source of information provided by more complex problem-solving tasks and the additional perspectives about students' achievement that can be gained. However, assessment results are often presented in aggregate form and the question arises about the precision of summary measures based on constructed-response tasks. More complex assessment tasks need longer assessment time, but many organizations and entities responsible for assessment are constrained by situations related to student burden levels and resources. Therefore, can assessments containing fewer, but potentially richer, questions be as precise as longer tests made up of multiple-choice questions?

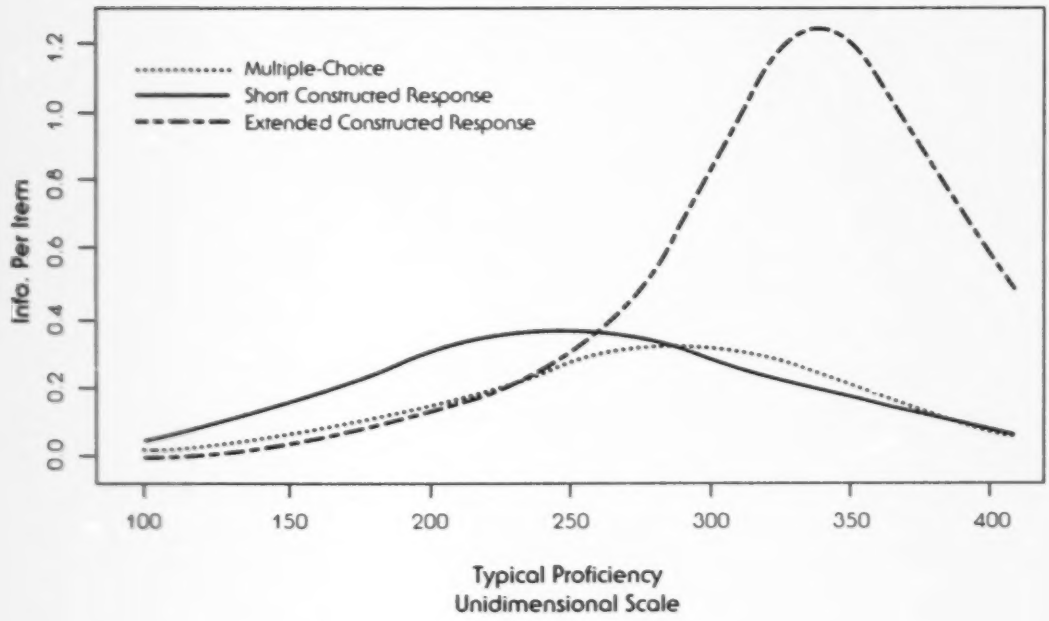
The primary way that NAEP summarizes student performance is on proficiency scales (ranging from 0 to 500) based on methods and concepts associated with item response theory (IRT). Although these results are not contained in this report, the scaling procedure forms the foundation of all the other NAEP reports about the 1992 mathematics assessment. All the questions in the assessment, including the constructed-response questions described herein, were used to develop the NAEP mathematics scales. For the extended-response questions, NAEP uses special partial-credit scaling techniques. Additionally, for grade 8, separate IRT-based measures of precision were calculated for the extended-response, regular constructed-response, and multiple-choice questions (see Procedural Appendix for further details). Each of these measures was then standardized so that results were expressed on a per-item basis, and comparisons were made between the various question types.[16]

The information functions, which show the per-item precision of measurement for varying degrees of mathematics proficiency, are shown in FIGURE 3.1. The information an item contributes to a test at a given proficiency level is given by the square of the ratio of the slope of the item characteristic curve at that point, to the standard error of measurement of observed scores at that proficiency level. The more highly discriminating an item, the more information it gives, and the more it contributes to shrinking the uncertainty of estimating proficiency from the observed responses. The maximum amount of

[16] John Mazzeo, Kentaro Yamamoto, and Edward Kulick, "Extended Constructed-Response Items in the 1992 NAEP: Psychometrically Speaking Were They Worth the Price?" Paper presented at the 1993 annual meeting of the National Council on Measurement in Education, Atlanta, GA.

information from the short constructed-response questions was obtained for students at about proficiency level 250.[17] For the multiple-choice items, the greatest amount of per-item information was obtained for students with proficiency levels near 285. For the extended-response questions, the greatest amount of per-item information was obtained for students with proficiencies of about 330. At the modal values, short constructed-response questions provided about 10 percent more information per question than did multiple-choice items. Extended-response questions provided about 3.5 times as much information as did the short constructed-response questions. Thus, at the proficiency levels where the question types measure most effectively, the extended-response tasks did provide substantially more information than their simpler, dichotomously scored counterparts.

Figure 3.1
Grade 8 Mathematics Per Item Information Functions



[17] Figure 3.1 contains a plot of the average information per item (on the y-axis) by proficiency level (on the x-axis) for the Grade 8 Mathematics multiple-choice, short constructed-response, and extended-response items based on unidimensional item parameters. The scale of the proficiency variable on the x-axis is centered on 250 and runs from about 100 to 400. The center of the scale roughly corresponds to the grade 8 mean score on the NAEP proficiency scale and each scale unit roughly corresponds to the standard deviation of grade 8 proficiency scores on the NAEP scale. However, somewhat different IRT scaling procedures were used for the analyses reported here than were used to produce the NAEP scales. Thus, the scales in Figure 3.1 are not directly comparable to the NAEP reported proficiency scale.

The expected per-item information for a randomly chosen eighth-grade student was .424 for the extended-response questions, .320 for the regular constructed-response questions, and .243 for the multiple-choice questions. The extended-response questions did not, however, provide better measurement across the entire range of proficiency. For proficiency values below 250, the extended-response tasks provided about the same information per item as the multiple-choice items. Both the multiple-choice and extended-response items provided somewhat less information than the short constructed-response items. Where extended-response tasks did show a marked superiority over the dichotomously scored items in terms of information provided was at the higher proficiency levels. The extended-response tasks provided considerably more information per-item than their binary counterparts above proficiency 300.

That the multiple-choice items provided less information than the short constructed-response questions for proficiency values below 250 may not be surprising given that the former are susceptible to guessing. The most plausible explanation for the surprisingly low amount of information provided by the extended-response tasks at low proficiency levels is the rather substantial level of difficulty exhibited by these items.

To briefly summarize, Figure 3.1 demonstrates the potential gain in information in IRT scaling that can be realized with extended-response tasks. In the proficiency range where they were effective, the extended-response tasks provided considerably more information than their simpler binary counterparts. At the modes of their respective information functions, the extended-response tasks provided about as much information as 4 multiple-choice items or 3 regular constructed-response questions. However, it is equally clear that the extended-response tasks were extremely difficult for students. For students in the lower half of the grade 8 proficiency distributions, the extended-response tasks provided little more in the way of information than multiple-choice items and less information than short constructed-response items. As a result, the expected per-task information for the extended-response tasks was not dramatically higher than the levels provided by the multiple-choice and short constructed-response items. However, as students become more familiar with performance tasks and activities as part of their school instructional programs, their performance probably will improve on these more complex assessment tasks.

Interaction Between Question Type and Curriculum and Instruction

As part of the 1992 mathematics assessment, NAEP collected a great deal of background information from students and their teachers. For example, teachers of fourth- and eighth-grade students who participated in the assessment were asked about the instructional emphasis placed on reasoning and communicating. The national results to these questions by performance on the three different question types are presented in TABLE 3.10. At grade 8, degree of emphasis on reasoning was related to performance regardless of question type. Some apparent evidence of this tendency can also be seen at grade 4, although the differences are not statistically significant. The results for emphasizing communication showed less clear relationships, although a pattern of non-statistically significant improved performance can be seen for the extended-response questions. The state-by-state results presented in TABLES 3.11 (extended-response), 3.12 (regular constructed-response), and 3.13 (multiple-choice) tend to mirror the national findings, although there are variations among the participants.

TABLE 3.10 National Average Percentages by Teachers' Reports on the Instructional Emphasis Placed on Reasoning and Communicating

Grades 4 and 8

	Developing Reasoning Ability to Solve Unique Problems			Learning How to Communicate Ideas in Mathematics Effectively		
	Heavy Emphasis	Moderate Emphasis	Little or No Emphasis	Heavy Emphasis	Moderate Emphasis	Little or No Emphasis
GRADE 4						
Satisfactory or Better on Extended-Response	17 (1.0)	16 (0.9)	14 (2.5)	17 (1.6)	16 (0.9)	15 (1.7)
Correct on Regular Constructed-Response	43 (1.0)	41 (0.8)	40 (1.8)	43 (1.3)	41 (0.7)	42 (1.2)
Correct on Multiple-Choice	50 (0.8)	49 (0.6)	48 (1.6)	49 (1.0)	49 (0.5)	50 (1.1)
GRADE 8						
Satisfactory on Better on Extended-Response	10 (0.8)	7 (0.6)	2 (0.9)	10 (0.9)	8 (0.7)	6 (1.4)
Correct on Regular Constructed-Response	57 (0.8)	52 (0.8)	44 (2.3)	56 (0.9)	53 (0.8)	53 (2.1)
Correct on Multiple-Choice	60 (0.7)	54 (0.7)	47 (1.7)	59 (0.7)	56 (0.6)	56 (1.6)

The standard errors of the estimated percentages appear in parentheses. It can be said with about 95 percent certainty for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details).

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment

TABLE 3.11 | **Average Percentage Satisfactory or Better Responses to Extended-Response Questions by Teacher's Reports on the Instructional Emphasis Placed on Reasoning and Communicating**

PUBLIC SCHOOLS	Grade 4 - 1992					
	Developing Reasoning Ability to Solve Unique Problems			Learning How to Communicate Ideas		
	Heavy Emphasis	Moderate Emphasis	Little or No Emphasis	Heavy Emphasis	Moderate Emphasis	Little or No Emphasis
NATION	16 (1.2)	16 (1.1)	13 (2.6)	16 (1.7)	16 (1.0)	15 (2.0)
Northeast	21 (2.3)	18 (2.3)	13 (6.0)	22 (2.4)	19 (2.8)	10 (4.8)
Southeast	13 (2.7)	13 (2.4)	4 (1.4)	12 (3.2)	13 (2.4)	7 (3.2)
Central	15 (2.1)	19 (2.6)	18 (4.7)	15 (4.2)	19 (1.8)	16 (3.2)
West	17 (2.1)	14 (1.5)	16 (6.6)	18 (3.6)	13 (1.6)	19 (3.0)
STATES						
Alabama	12 (1.3)	11 (1.2)	9 (3.3)	11 (1.2)	12 (1.3)	11 (2.5)
Arizona	15 (0.9)	12 (1.1)	15 (2.4)	14 (1.1)	13 (1.1)	15 (2.6)
Arkansas	11 (1.1)	11 (0.8)	11 (2.9)	9 (1.3)	11 (0.7)	12 (2.2)
California	13 (1.0)	11 (1.1)	4 (2.2)	12 (1.3)	11 (0.9)	10 (2.9)
Colorado	16 (1.0)	16 (1.3)	17 (4.5)	17 (1.1)	15 (1.2)	17 (2.5)
Connecticut	25 (1.5)	20 (1.5)	13 (5.6)	24 (1.7)	23 (1.5)	17 (2.4)
Delaware	16 (1.1)	14 (1.3)	14 (2.0)	18 (1.0)	13 (1.3)	15 (2.0)
Dist. Columbia	9 (1.2)	5 (1.0)	4 (2.6)	8 (1.0)	7 (1.4)	3 (1.8)
Florida	14 (1.2)	10 (1.1)	7 (2.6)	13 (1.1)	12 (1.2)	8 (2.4)
Georgia	17 (1.2)	13 (1.2)	14 (4.1)	15 (1.1)	15 (1.3)	14 (3.5)
Hawaii	14 (1.4)	11 (1.2)	15 (3.4)	13 (1.6)	13 (1.2)	10 (2.3)
Idaho	16 (1.2)	16 (1.3)	15 (4.2)	17 (1.6)	15 (1.1)	16 (1.8)
Indiana	18 (1.5)	14 (1.4)	10 (2.2)	14 (1.5)	16 (1.6)	16 (2.4)
Iowa	25 (1.3)	22 (1.5)	16 (4.2)	22 (1.9)	24 (1.2)	20 (2.4)
Kentucky	15 (1.2)	13 (1.5)	8 (3.0)	15 (1.3)	14 (1.5)	13 (2.9)
Louisiana	9 (0.9)	7 (1.1)	6 (1.7)	9 (1.0)	8 (1.1)	4 (2.5)
Maine	23 (1.8)	22 (1.6)	13 (5.0)	23 (2.0)	23 (1.7)	18 (2.4)
Maryland	18 (1.4)	16 (1.7)	16 (7.9)	17 (1.4)	19 (1.3)	19 (4.3)
Massachusetts	19 (1.8)	19 (1.4)	12 (2.9)	21 (2.2)	18 (1.6)	16 (1.8)
Michigan	15 (1.1)	17 (1.6)	11 (6.6)	15 (1.6)	17 (1.3)	11 (2.3)
Minnesota	20 (1.1)	19 (1.6)	19 (3.5)	18 (1.7)	20 (1.3)	22 (2.8)
Mississippi	8 (1.1)	7 (1.0)	6 (3.0)	6 (0.9)	9 (1.1)	6 (2.7)
Missouri	19 (1.4)	17 (1.3)	16 (2.5)	18 (1.6)	18 (1.3)	18 (2.3)
Nebraska	19 (1.5)	18 (1.3)	15 (5.2)	17 (1.6)	20 (1.4)	15 (2.4)
New Hampshire	23 (1.4)	20 (1.4)	19 (7.6)	25 (2.0)	19 (1.4)	19 (2.6)
New Jersey	20 (1.4)	17 (1.7)	15 (3.4)	20 (1.3)	18 (1.7)	14 (3.3)
New Mexico	14 (1.6)	11 (1.1)	11 (2.7)	12 (1.7)	13 (1.1)	10 (2.7)
New York	14 (1.3)	15 (1.5)	17 (4.6)	14 (1.6)	14 (1.4)	21 (2.7)
North Carolina	14 (1.0)	13 (1.2)	9 (2.3)	13 (1.4)	13 (0.9)	15 (2.4)
North Dakota	22 (1.5)	20 (1.1)	10 (3.4)	21 (1.8)	20 (1.1)	18 (2.8)
Ohio	19 (1.2)	16 (1.4)	9 (2.6)	18 (1.5)	17 (1.2)	10 (2.3)
Oklahoma	16 (1.1)	16 (1.3)	11 (3.0)	16 (1.4)	16 (1.3)	15 (2.5)
Pennsylvania	21 (1.3)	18 (1.4)	11 (3.5)	21 (1.6)	18 (1.2)	16 (3.2)
Rhode Island	17 (1.6)	14 (1.5)	14 (3.5)	15 (2.1)	16 (1.6)	14 (2.0)
South Carolina	12 (0.8)	11 (1.1)	9 (2.8)	13 (1.0)	10 (0.9)	10 (2.5)
Tennessee	14 (1.3)	13 (1.1)	10 (2.6)	13 (1.1)	14 (1.0)	13 (2.4)
Texas	17 (1.3)	13 (1.3)	17 (6.4)	16 (1.2)	16 (1.6)	8 (2.1)
Utah	16 (1.1)	16 (1.2)	13 (1.9)	17 (1.4)	15 (1.1)	17 (1.4)
Virginia	19 (1.6)	16 (1.5)	15 (2.7)	15 (1.4)	20 (1.7)	15 (2.2)
West Virginia	14 (1.2)	11 (0.9)	11 (2.3)	13 (1.3)	11 (0.9)	13 (1.7)
Wisconsin	23 (1.6)	18 (1.6)	20 (4.0)	23 (1.7)	19 (1.3)	21 (2.5)
Wyoming	19 (1.1)	18 (1.2)	10 (2.6)	20 (1.6)	18 (1.3)	16 (1.8)
TERRITORY						
Guam	6 (0.8)	8 (0.9)	4 (1.1)	7 (0.7)	6 (0.9)	7 (1.4)

The standard errors of the estimated percentages appear in parentheses. It can be said with 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details).

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment.

TABLE 3.11 | **Average Percentage Satisfactory or Better Responses to Extended-Response Questions by Teacher's Reports on the Instructional Emphasis Placed on Reasoning and Communicating (continued)**

PUBLIC SCHOOLS	Grade 8 - 1992					
	Developing Reasoning Ability to Solve Unique Problems			Learning How to Communicate Ideas		
	Heavy Emphasis	Moderate Emphasis	Little or No Emphasis	Heavy Emphasis	Moderate Emphasis	Little or No Emphasis
NATION	13 (0.8)	10 (0.8)	8 (1.8)	13 (0.9)	11 (0.8)	10 (1.6)
Northeast	17 (1.9)	13 (2.6)	3 (1.9)	18 (2.2)	12 (1.8)	19 (5.5)
Southeast	11 (1.3)	9 (1.1)	9 (2.1)	10 (1.3)	11 (1.3)	8 (3.9)
Central	11 (1.2)	10 (1.7)	9 (2.9)	9 (1.1)	11 (2.2)	11 (1.8)
West	13 (1.6)	8 (1.2)	11 (4.3)	13 (1.6)	10 (1.2)	5 (1.6)
STATES						
Alabama	9 (0.9)	6 (0.7)	4 (1.6)	8 (0.9)	6 (0.8)	3 (1.1)
Arizona	12 (1.1)	7 (0.9)	7 (2.2)	11 (1.3)	9 (1.3)	7 (1.2)
Arkansas	10 (1.0)	8 (0.7)	6 (1.3)	8 (0.8)	8 (0.8)	9 (1.4)
California	13 (1.3)	11 (0.9)	10 (2.1)	12 (1.9)	13 (0.9)	9 (1.5)
Colorado	15 (0.8)	10 (1.1)	9 (2.6)	14 (1.0)	12 (0.9)	8 (1.4)
Connecticut	16 (1.1)	13 (1.1)	10 (2.0)	15 (1.2)	14 (1.1)	11 (1.6)
Delaware	10 (0.9)	9 (1.0)	8 (2.3)	9 (0.9)	10 (1.0)	9 (2.5)
Dist. Columbia	13 (1.1)	16 (3.0)	9 (3.1)	12 (1.2)	17 (1.9)	11(14.2)
Florida	11 (0.8)	10 (0.8)	8 (1.7)	11 (0.9)	10 (0.9)	9 (1.7)
Georgia	11 (1.0)	8 (0.8)	7 (1.6)	10 (0.9)	9 (0.7)	6 (1.6)
Hawaii	11 (1.0)	8 (0.8)	7 (1.3)	11 (1.1)	8 (0.7)	7 (1.2)
Idaho	13 (1.1)	9 (0.8)	9 (3.5)	12 (1.0)	11 (1.1)	9 (1.7)
Indiana	14 (1.0)	8 (0.9)	10 (2.4)	13 (1.3)	10 (0.9)	9 (1.8)
Iowa	17 (1.2)	13 (0.9)	7 (1.5)	15 (1.1)	15 (1.1)	13 (2.8)
Kentucky	12 (1.0)	7 (0.8)	4 (1.3)	12 (1.0)	7 (0.7)	5 (1.0)
Louisiana	8 (0.9)	8 (0.8)	5 (1.3)	9 (0.8)	7 (0.7)	7 (1.8)
Maine	15 (1.2)	10 (1.1)	5 (1.4)	17 (1.6)	10 (0.9)	8 (2.7)
Maryland	17 (1.3)	9 (1.0)	9 (2.2)	16 (1.5)	10 (1.0)	12 (2.3)
Massachusetts	16 (1.3)	11 (1.2)	8 (1.6)	15 (1.3)	12 (1.2)	7 (1.3)
Michigan	15 (1.1)	10 (1.1)	5 (2.9)	14 (1.2)	11 (0.9)	11 (2.7)
Minnesota	17 (1.2)	12 (1.1)	8 (2.4)	18 (1.5)	12 (1.1)	12 (1.7)
Mississippi	9 (0.9)	7 (1.1)	6 (1.5)	8 (0.9)	7 (1.1)	7 (1.7)
Missouri	13 (1.3)	9 (0.8)	7 (1.8)	12 (1.4)	10 (0.8)	8 (1.3)
Nebraska	15 (1.2)	11 (0.9)	5 (2.0)	14 (1.4)	12 (1.0)	10 (1.2)
New Hampshire	15 (1.1)	12 (1.1)	12 (4.0)	15 (1.3)	12 (0.9)	12 (2.5)
New Jersey	15 (1.0)	11 (1.2)	7 (2.6)	16 (1.1)	11 (1.1)	8 (2.8)
New Mexico	10 (0.8)	7 (0.7)	6 (1.3)	10 (0.8)	8 (0.7)	6 (1.3)
New York	14 (1.3)	12 (1.1)	6 (3.0)	15 (1.0)	12 (1.0)	10 (1.5)
North Carolina	11 (0.9)	7 (0.8)	9 (2.5)	10 (0.9)	8 (0.8)	9 (2.1)
North Dakota	16 (1.3)	11 (1.2)	13 (2.8)	14 (1.5)	13 (1.1)	13 (2.4)
Ohio	14 (1.3)	10 (1.1)	8 (1.9)	14 (1.4)	11 (1.0)	10 (1.5)
Oklahoma	11 (1.1)	9 (1.1)	8 (1.0)	10 (1.2)	9 (1.2)	10 (1.7)
Pennsylvania	15 (1.3)	8 (1.2)	12 (2.6)	14 (1.3)	11 (1.2)	11 (2.9)
Rhode Island	14 (1.0)	8 (0.9)	7 (1.3)	14 (1.1)	8 (0.9)	9 (1.5)
South Carolina	11 (0.9)	8 (0.8)	7 (1.9)	10 (0.9)	9 (1.0)	7 (2.1)
Tennessee	9 (0.9)	7 (0.7)	3 (0.9)	8 (0.8)	7 (0.6)	5 (1.8)
Texas	13 (1.1)	11 (1.2)	7 (1.7)	13 (1.1)	11 (1.3)	10 (1.6)
Utah	14 (1.1)	10 (0.9)	11 (2.4)	15 (1.5)	10 (0.9)	12 (1.5)
Virginia	14 (1.0)	10 (0.8)	10 (1.4)	14 (1.1)	11 (0.7)	9 (2.1)
West Virginia	9 (0.7)	7 (0.6)	6 (1.2)	9 (1.0)	8 (0.7)	5 (1.0)
Wisconsin	15 (1.3)	11 (1.4)	9 (1.8)	15 (1.9)	12 (1.2)	11 (3.5)
Wyoming	13 (1.2)	9 (0.8)	6 (1.2)	13 (1.4)	9 (0.9)	8 (1.2)
TERRITORIES						
Guam	13 (1.6)	11 (1.1)	21 (2.3)	14 (1.5)	11 (1.2)	15 (1.5)
Virgin Islands	13 (1.0)	17 (1.3)	30(30.2)	14 (0.9)	15 (1.8)	14 (2.5)

TABLE 3.12 | **Overall Average Percentage Correct for Regular Constructed-Response Questions by Teacher's Reports on the Instructional Emphasis Placed on Reasoning and Communicating**

PUBLIC SCHOOLS	Grade 4 - 1992					
	Developing Reasoning Ability to Solve Unique Problems			Learning How to Communicate Ideas		
	Heavy Emphasis	Moderate Emphasis	Little or No Emphasis	Heavy Emphasis	Moderate Emphasis	Little or No Emphasis
NATION	43 (1.1)	40 (1.0)	39 (2.2)	42 (1.5)	40 (0.8)	41 (1.5)
Northeast	46 (2.3)	40 (2.5)	40 (4.5)	46 (2.4)	43 (2.3)	32 (3.1)
Southeast	36 (2.4)	34 (2.3)	31 (4.3)	36 (3.5)	34 (1.6)	35 (2.5)
Central	45 (2.1)	44 (1.4)	46 (1.9)	45 (2.4)	45 (1.5)	46 (1.9)
West	44 (2.5)	40 (1.3)	37 (9.7)	46 (2.9)	39 (1.0)	41 (2.8)
STATES						
Alabama	35 (1.4)	33 (1.3)	33 (2.8)	34 (1.3)	35 (1.5)	34 (2.6)
Arizona	41 (0.9)	40 (1.0)	40 (2.2)	40 (1.1)	40 (0.9)	41 (2.3)
Arkansas	37 (1.2)	36 (0.7)	30 (2.7)	36 (1.6)	36 (0.7)	35 (2.0)
California	37 (1.3)	37 (1.2)	29 (2.1)	38 (1.5)	37 (1.1)	28 (2.3)
Colorado	45 (1.0)	43 (1.0)	43 (4.0)	46 (1.1)	43 (1.1)	43 (2.1)
Connecticut	51 (1.1)	45 (1.2)	36 (6.6)	50 (1.4)	49 (1.0)	41 (2.1)
Delaware	44 (1.0)	39 (0.7)	37 (1.6)	43 (0.8)	39 (1.0)	38 (1.2)
Dist. Columbia	29 (0.5)	25 (0.6)	25 (1.3)	29 (0.6)	25 (0.8)	27 (1.1)
Florida	41 (0.8)	36 (1.2)	33 (2.7)	39 (1.1)	39 (1.1)	35 (2.9)
Georgia	41 (1.4)	37 (1.5)	36 (4.2)	40 (1.3)	39 (1.3)	35 (2.6)
Hawaii	43 (1.3)	38 (1.0)	35 (2.7)	41 (1.3)	40 (1.0)	36 (1.9)
Idaho	44 (1.0)	43 (0.9)	43 (3.0)	45 (1.2)	44 (0.8)	39 (1.6)
Indiana	44 (1.4)	42 (0.8)	39 (2.8)	42 (1.7)	43 (1.0)	41 (1.5)
Iowa	50 (0.9)	49 (0.9)	43 (2.7)	49 (1.3)	50 (0.9)	49 (1.7)
Kentucky	39 (0.8)	37 (1.2)	37 (5.7)	38 (1.0)	38 (1.1)	40 (3.6)
Louisiana	34 (1.0)	32 (1.4)	27 (2.4)	33 (1.2)	33 (1.2)	29 (2.7)
Maine	52 (1.1)	50 (1.0)	45 (2.7)	52 (1.4)	51 (0.9)	48 (1.7)
Maryland	44 (1.4)	41 (1.5)	32 (4.6)	42 (1.2)	44 (1.3)	41 (3.6)
Massachusetts	49 (1.3)	47 (1.4)	37 (2.7)	48 (1.9)	48 (1.3)	44 (2.7)
Michigan	42 (1.4)	45 (1.5)	34 (5.4)	41 (1.9)	45 (1.3)	37 (2.0)
Minnesota	50 (1.1)	47 (1.4)	42 (2.4)	49 (1.5)	47 (1.1)	50 (1.5)
Mississippi	29 (1.1)	31 (1.0)	25 (1.8)	28 (1.0)	31 (1.1)	31 (2.6)
Missouri	46 (1.3)	43 (1.2)	41 (2.3)	45 (1.6)	45 (1.2)	41 (1.6)
Nebraska	48 (1.3)	46 (1.2)	44 (3.4)	49 (1.7)	47 (1.0)	42 (2.4)
New Hampshire	52 (1.2)	48 (1.2)	44 (2.0)	53 (1.5)	49 (1.1)	47 (2.1)
New Jersey	49 (1.4)	44 (1.3)	42 (2.6)	49 (1.5)	45 (1.4)	43 (2.9)
New Mexico	42 (1.5)	37 (0.9)	37 (3.1)	40 (1.5)	38 (1.2)	37 (2.1)
New York	42 (1.2)	41 (1.2)	43 (4.7)	41 (1.3)	42 (1.2)	42 (2.4)
North Carolina	39 (0.9)	37 (1.1)	36 (3.0)	38 (1.0)	38 (0.9)	41 (1.9)
North Dakota	49 (1.0)	48 (0.9)	43 (1.7)	47 (1.1)	48 (0.8)	48 (2.1)
Ohio	44 (1.0)	40 (1.2)	30 (3.4)	43 (1.4)	41 (1.1)	38 (2.3)
Oklahoma	42 (1.1)	42 (1.0)	44 (2.0)	42 (1.0)	42 (1.1)	41 (1.7)
Pennsylvania	48 (1.2)	43 (1.4)	39 (3.0)	48 (1.4)	44 (1.2)	41 (3.2)
Rhode Island	41 (1.9)	38 (1.0)	36 (3.1)	38 (2.2)	40 (1.4)	38 (2.0)
South Carolina	38 (1.0)	37 (1.2)	33 (4.2)	38 (1.1)	36 (1.1)	37 (2.4)
Tennessee	37 (1.5)	36 (0.9)	28 (1.9)	36 (1.4)	36 (1.0)	33 (1.9)
Texas	42 (1.0)	39 (1.8)	36 (3.3)	42 (1.4)	40 (1.3)	33 (3.1)
Utah	46 (1.0)	44 (1.0)	39 (2.4)	48 (1.3)	43 (0.9)	41 (1.7)
Virginia	45 (1.4)	40 (1.4)	43 (3.4)	42 (1.3)	44 (1.3)	42 (2.6)
West Virginia	40 (1.2)	39 (1.0)	33 (2.2)	39 (1.3)	38 (0.8)	39 (1.9)
Wisconsin	50 (0.9)	46 (1.1)	42 (3.2)	49 (1.1)	48 (1.1)	47 (1.7)
Wyoming	47 (0.9)	45 (0.7)	38 (2.0)	47 (1.3)	45 (0.8)	47 (1.4)
TERRITORY						
Guam	28 (0.8)	29 (0.6)	21 (1.3)	31 (0.7)	25 (0.7)	27 (1.0)

The standard errors of the estimated percentages appear in parentheses. It can be said with 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details).

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment.

TABLE 3.12 | **Overall Average Percentage Correct for Regular Constructed Response Questions by Teacher's Reports on the Instructional Emphasis Placed on Reasoning and Communicating (continued)**

PUBLIC SCHOOLS	Grade 8 - 1992					
	Developing Reasoning Ability to Solve Unique Problems			Learning How to Communicate Ideas		
	Heavy Emphasis	Moderate Emphasis	Little or No Emphasis	Heavy Emphasis	Moderate Emphasis	Little or No Emphasis
NATION	56 (0.9)	51 (0.8)	43 (2.5)	55 (1.0)	52 (0.8)	52 (2.4)
Northeast	55 (2.1)	51 (3.4)	53 (4.0)	57 (2.8)	52 (2.1)	47 (5.8)
Southeast	52 (1.4)	48 (1.1)	44 (4.3)	51 (1.6)	49 (1.3)	48 (5.8)
Central	62 (2.1)	57 (1.4)	46 (5.5)	60 (1.7)	57 (1.8)	59 (4.4)
West	57 (1.3)	50 (1.1)	34 (4.1)	55 (1.4)	52 (1.4)	43 (5.0)
STATES						
Alabama	49 (1.9)	45 (1.1)	38 (2.8)	48 (1.7)	45 (1.3)	40 (2.7)
Arizona	56 (0.8)	51 (1.1)	41 (2.6)	55 (1.0)	54 (1.1)	46 (2.3)
Arkansas	54 (1.4)	47 (1.0)	42 (2.2)	51 (1.4)	48 (1.1)	44 (1.5)
California	56 (1.2)	49 (1.2)	40 (4.4)	54 (1.3)	52 (1.2)	46 (2.3)
Colorado	61 (0.7)	55 (1.1)	51 (2.7)	59 (1.2)	57 (0.9)	56 (1.7)
Connecticut	62 (1.2)	56 (1.1)	48 (3.1)	61 (1.1)	57 (1.3)	50 (2.8)
Delaware	56 (0.8)	51 (0.9)	42 (2.1)	55 (1.0)	52 (0.8)	49 (1.7)
Dist. Columbia	37 (0.8)	39 (1.3)	33 (2.4)	36 (0.8)	39 (1.2)	22 (2.3)
Florida	55 (1.0)	48 (1.2)	39 (2.2)	54 (1.1)	48 (1.0)	46 (2.4)
Georgia	52 (1.1)	46 (1.3)	42 (2.0)	50 (1.0)	48 (1.4)	45 (2.0)
Hawaii	58 (0.9)	48 (0.8)	39 (1.0)	57 (0.9)	48 (0.7)	41 (1.1)
Idaho	62 (0.8)	57 (0.7)	54 (3.7)	60 (0.8)	60 (0.8)	56 (1.0)
Indiana	61 (0.9)	53 (1.1)	47 (2.6)	60 (1.1)	55 (1.0)	52 (2.1)
Iowa	66 (0.9)	61 (1.1)	56 (2.0)	64 (1.0)	62 (0.8)	60 (3.0)
Kentucky	56 (1.1)	49 (1.1)	37 (2.4)	56 (1.2)	50 (1.0)	46 (2.6)
Louisiana	48 (1.7)	42 (1.0)	38 (2.8)	46 (1.6)	44 (1.2)	43 (2.7)
Maine	65 (1.0)	59 (0.9)	56 (2.5)	65 (1.1)	61 (0.6)	54 (1.9)
Maryland	59 (1.3)	49 (1.3)	44 (2.5)	56 (1.4)	52 (1.3)	50 (4.0)
Massachusetts	63 (0.9)	54 (1.3)	48 (2.1)	62 (1.4)	56 (1.2)	52 (2.3)
Michigan	56 (1.5)	54 (1.3)	49 (3.9)	55 (1.6)	56 (1.2)	56 (2.5)
Minnesota	66 (0.8)	60 (0.8)	51 (3.5)	65 (1.1)	61 (0.9)	60 (0.9)
Mississippi	44 (1.1)	41 (1.3)	38 (2.4)	43 (1.0)	42 (1.4)	41 (3.7)
Missouri	62 (1.0)	54 (0.9)	53 (2.0)	60 (1.2)	56 (0.9)	56 (1.5)
Nebraska	64 (1.0)	58 (0.8)	55 (3.6)	62 (1.0)	60 (0.9)	59 (1.9)
New Hampshire	64 (0.9)	59 (0.8)	57 (3.7)	64 (0.8)	60 (0.8)	54 (1.9)
New Jersey	58 (1.2)	55 (1.7)	48 (3.3)	58 (1.3)	56 (1.3)	56 (3.6)
New Mexico	55 (1.0)	48 (0.9)	44 (1.4)	54 (0.8)	49 (1.0)	46 (1.3)
New York	59 (1.6)	52 (1.6)	40 (3.0)	58 (1.7)	52 (1.7)	53 (3.4)
North Carolina	54 (1.0)	46 (1.1)	45 (2.7)	52 (1.2)	47 (1.0)	51 (1.7)
North Dakota	64 (0.7)	63 (1.0)	59 (3.7)	64 (0.9)	63 (0.8)	63 (2.5)
Ohio	59 (1.3)	55 (1.1)	48 (3.2)	59 (1.7)	54 (1.2)	55 (2.0)
Oklahoma	60 (1.3)	53 (0.9)	47 (1.6)	57 (1.2)	55 (0.9)	50 (2.1)
Pennsylvania	60 (1.1)	54 (1.0)	52 (3.6)	59 (1.2)	55 (1.1)	53 (3.7)
Rhode Island	57 (0.5)	51 (0.8)	44 (1.9)	56 (0.6)	52 (0.7)	50 (1.9)
South Carolina	54 (1.0)	47 (0.9)	41 (3.0)	51 (1.0)	49 (1.1)	49 (4.0)
Tennessee	53 (1.1)	47 (1.0)	44 (3.4)	51 (1.2)	49 (0.9)	45 (3.1)
Texas	56 (1.0)	50 (1.4)	40 (2.5)	54 (1.3)	53 (1.1)	44 (2.1)
Utah	62 (0.8)	57 (1.1)	53 (1.8)	61 (1.0)	58 (1.0)	55 (1.4)
Virginia	59 (1.2)	52 (1.2)	44 (1.9)	58 (1.4)	53 (1.1)	45 (2.6)
West Virginia	55 (1.0)	47 (1.1)	40 (1.6)	55 (1.2)	48 (0.9)	45 (1.7)
Wisconsin	63 (1.0)	59 (0.9)	51 (3.1)	63 (1.5)	60 (0.8)	56 (2.9)
Wyoming	62 (0.8)	58 (0.8)	53 (1.3)	63 (0.7)	57 (0.7)	57 (1.4)
TERRITORIES						
Guam	46 (1.1)	35 (1.1)	29 (1.5)	43 (0.9)	39 (1.1)	27 (1.3)
Virgin Islands						

TABLE 3.13 | **Overall Average Percentage Correct for Multiple-Choice Questions by Teacher's Reports on the Instructional Emphasis Placed on Reasoning and Communicating**

PUBLIC SCHOOLS	Grade 4 - 1992					
	Developing Reasoning Ability to Solve Unique Problems			Learning How to Communicate Ideas		
	Heavy Emphasis	Moderate Emphasis	Little or No Emphasis	Heavy Emphasis	Moderate Emphasis	Little or No Emphasis
NATION	49 (0.8)	48 (0.7)	48 (2.0)	49 (1.1)	48 (0.6)	49 (1.3)
Northeast	53 (1.6)	47 (1.8)	43 (4.4)	53 (1.9)	50 (1.5)	38 (1.5)
Southeast	45 (1.2)	44 (1.4)	42 (3.0)	44 (1.9)	44 (1.3)	46 (2.4)
Central	51 (1.8)	51 (1.4)	53 (2.6)	51 (2.0)	51 (1.5)	52 (1.6)
West	49 (2.0)	49 (1.0)	48 (4.3)	51 (2.4)	48 (0.8)	49 (2.6)
STATES						
Alabama	46 (0.9)	44 (1.0)	44 (2.8)	45 (0.9)	46 (1.2)	45 (2.2)
Arizona	48 (0.6)	47 (0.7)	46 (1.9)	48 (0.8)	47 (0.7)	47 (1.6)
Arkansas	46 (0.7)	47 (0.5)	43 (1.8)	46 (1.1)	47 (0.5)	45 (1.4)
California	46 (0.9)	46 (0.9)	39 (2.2)	46 (1.1)	46 (0.8)	40 (1.8)
Colorado	51 (0.7)	49 (0.8)	50 (3.7)	51 (0.8)	49 (0.8)	49 (1.9)
Connecticut	56 (0.8)	52 (0.9)	42 (4.3)	54 (1.0)	54 (0.8)	50 (1.8)
Delaware	53 (0.5)	47 (0.6)	47 (1.3)	52 (0.5)	48 (0.9)	47 (1.2)
Dist. Columbia	39 (0.5)	37 (0.5)	39 (2.0)	39 (0.5)	37 (0.6)	37 (1.0)
Florida	49 (0.7)	46 (0.8)	45 (1.9)	48 (0.8)	48 (0.8)	45 (2.1)
Georgia	50 (0.9)	47 (1.2)	47 (2.4)	49 (1.0)	49 (0.9)	44 (1.9)
Hawaii	50 (1.0)	46 (0.7)	44 (2.0)	49 (1.1)	48 (0.7)	44 (1.1)
Idaho	51 (0.7)	49 (0.6)	49 (1.7)	51 (0.8)	50 (0.6)	47 (1.2)
Indiana	51 (1.0)	49 (0.6)	46 (2.0)	50 (1.2)	50 (0.8)	49 (1.2)
Iowa	55 (0.7)	54 (0.7)	49 (2.1)	54 (1.0)	55 (0.7)	52 (0.8)
Kentucky	48 (0.7)	47 (0.8)	45 (2.9)	47 (0.7)	48 (0.8)	47 (2.4)
Louisiana	44 (0.7)	43 (1.1)	40 (1.2)	42 (0.8)	44 (0.9)	40 (1.4)
Maine	57 (0.8)	54 (0.9)	53 (2.2)	57 (1.1)	55 (0.7)	55 (1.5)
Maryland	50 (1.0)	49 (1.1)	41 (3.6)	48 (1.0)	51 (1.0)	48 (2.5)
Massachusetts	54 (0.9)	53 (1.0)	47 (2.4)	54 (1.3)	53 (0.9)	51 (1.8)
Michigan	50 (1.0)	51 (1.2)	46 (3.2)	49 (1.3)	52 (1.0)	46 (1.5)
Minnesota	55 (0.8)	54 (0.8)	50 (2.0)	54 (1.1)	54 (0.6)	56 (1.4)
Mississippi	42 (0.8)	43 (0.7)	38 (2.0)	41 (0.8)	43 (0.8)	44 (2.7)
Missouri	52 (1.0)	50 (0.9)	48 (1.8)	50 (1.1)	52 (0.9)	50 (1.3)
Nebraska	54 (0.8)	52 (0.9)	49 (2.2)	54 (1.2)	53 (0.9)	52 (2.0)
New Hampshire	56 (0.9)	53 (0.9)	51 (1.3)	56 (1.1)	53 (0.8)	53 (1.9)
New Jersey	56 (1.0)	52 (1.0)	49 (1.5)	55 (1.0)	53 (1.1)	50 (2.6)
New Mexico	48 (1.1)	45 (0.7)	43 (1.3)	47 (0.8)	46 (1.0)	45 (1.2)
New York	50 (0.9)	49 (0.8)	52 (3.1)	49 (1.0)	50 (0.7)	50 (1.5)
North Carolina	48 (0.6)	47 (0.9)	46 (2.2)	47 (0.8)	47 (0.7)	48 (1.4)
North Dakota	55 (0.8)	54 (0.9)	49 (1.7)	54 (0.9)	54 (0.9)	54 (1.1)
Ohio	51 (0.8)	48 (0.9)	42 (2.2)	50 (1.0)	49 (0.8)	47 (1.9)
Oklahoma	50 (0.8)	49 (0.8)	50 (2.2)	49 (0.9)	50 (0.7)	49 (1.3)
Pennsylvania	54 (1.0)	51 (1.0)	48 (2.4)	54 (1.1)	51 (0.9)	49 (2.0)
Rhode Island	49 (1.2)	47 (0.8)	45 (2.6)	48 (1.5)	49 (1.0)	47 (1.4)
South Carolina	47 (0.7)	46 (0.9)	46 (3.7)	48 (0.7)	46 (0.9)	47 (1.6)
Tennessee	46 (1.0)	46 (0.7)	43 (1.5)	45 (1.0)	46 (0.8)	46 (1.3)
Texas	51 (0.7)	49 (1.0)	50 (2.9)	51 (1.0)	50 (0.9)	46 (1.7)
Utah	54 (0.7)	51 (0.6)	48 (2.2)	55 (0.9)	51 (0.6)	50 (1.3)
Virginia	51 (1.1)	49 (1.0)	48 (2.3)	49 (0.9)	51 (1.0)	49 (1.4)
West Virginia	48 (1.0)	47 (0.7)	44 (1.7)	48 (1.1)	47 (0.6)	47 (1.3)
Wisconsin	56 (0.8)	53 (0.9)	49 (2.6)	55 (1.0)	54 (0.7)	55 (1.5)
Wyoming	52 (0.8)	51 (0.5)	48 (2.0)	52 (0.9)	51 (0.7)	53 (0.9)
TERRITORY						
Guam	40 (0.6)	40 (0.4)	35 (0.8)	41 (0.5)	38 (0.5)	38 (0.8)

The standard errors of the estimated percentages appear in parentheses. It can be said with 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details).

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment.

TABLE 3.13 | **Overall Average Percentage Correct for Multiple-Choice Questions by Teacher's Reports on the Instructional Emphasis Placed on Reasoning and Communicating (continued)**

PUBLIC SCHOOLS	Grade 8 - 1992					
	Developing Reasoning Ability to Solve Unique Problems			Learning How to Communicate Ideas		
	Heavy Emphasis	Moderate Emphasis	Little or No Emphasis	Heavy Emphasis	Moderate Emphasis	Little or No Emphasis
NATION	60 (0.7)	54 (0.7)	46 (1.7)	58 (0.8)	55 (0.6)	55 (1.8)
Northeast	59 (1.5)	54 (2.9)	50 (3.4)	60 (2.1)	55 (1.5)	50 (4.7)
Southeast	57 (1.5)	51 (1.2)	49 (2.6)	56 (1.7)	53 (0.5)	51 (4.0)
Central	64 (1.2)	58 (1.3)	45 (4.0)	62 (1.2)	58 (1.8)	60 (2.9)
West	60 (1.2)	53 (1.0)	42 (3.3)	58 (1.3)	56 (1.3)	48 (1.9)
STATES						
Alabama	54 (1.3)	49 (1.0)	42 (1.4)	52 (1.1)	50 (1.1)	46 (2.1)
Arizona	59 (0.7)	53 (1.0)	45 (1.7)	58 (0.9)	55 (0.7)	49 (1.7)
Arkansas	57 (1.0)	51 (0.7)	46 (1.5)	55 (1.1)	51 (0.8)	48 (1.2)
California	58 (1.2)	52 (1.1)	46 (2.4)	56 (1.2)	55 (1.1)	49 (1.4)
Colorado	61 (0.7)	56 (0.9)	53 (1.9)	59 (1.0)	58 (0.8)	57 (1.2)
Connecticut	63 (1.0)	58 (1.1)	51 (2.3)	62 (1.1)	59 (1.1)	54 (2.3)
Delaware	58 (0.7)	52 (0.7)	45 (1.9)	56 (0.8)	55 (0.6)	48 (1.6)
Dist. Columbia	44 (0.5)	45 (1.1)	41 (1.5)	44 (0.6)	45 (0.9)	33 (9.1)
Florida	58 (0.7)	51 (1.0)	43 (2.0)	56 (0.9)	52 (0.8)	51 (2.3)
Georgia	56 (0.9)	50 (0.9)	47 (1.8)	55 (0.7)	52 (1.1)	49 (1.4)
Hawaii	59 (0.7)	51 (0.6)	43 (1.0)	58 (0.6)	51 (0.6)	47 (1.0)
Idaho	63 (0.7)	58 (0.6)	54 (2.6)	61 (0.6)	61 (0.7)	57 (1.1)
Indiana	62 (0.9)	55 (0.9)	50 (2.5)	61 (1.1)	57 (0.8)	54 (1.9)
Iowa	67 (0.8)	62 (0.9)	57 (2.3)	66 (0.9)	63 (0.8)	61 (2.9)
Kentucky	58 (0.9)	52 (0.7)	44 (1.7)	58 (1.0)	52 (0.7)	49 (1.9)
Louisiana	52 (1.3)	48 (0.8)	46 (2.1)	50 (1.1)	49 (0.9)	48 (1.9)
Maine	65 (0.9)	58 (0.8)	55 (2.4)	65 (1.1)	60 (0.7)	55 (2.0)
Maryland	61 (1.2)	53 (1.0)	49 (1.9)	59 (1.2)	55 (1.1)	54 (3.1)
Massachusetts	63 (0.9)	56 (1.0)	49 (1.7)	62 (1.1)	57 (0.9)	53 (1.7)
Michigan	58 (1.2)	55 (1.1)	54 (3.3)	57 (1.3)	57 (1.1)	57 (1.8)
Minnesota	68 (0.8)	61 (0.8)	53 (2.7)	67 (1.1)	63 (0.8)	60 (1.0)
Mississippi	50 (0.9)	47 (1.0)	44 (1.6)	49 (0.8)	47 (0.9)	47 (1.9)
Missouri	62 (1.0)	56 (0.7)	54 (2.0)	61 (1.1)	57 (0.7)	57 (1.3)
Nebraska	65 (0.9)	60 (0.9)	55 (2.7)	64 (0.9)	61 (0.7)	61 (1.8)
New Hampshire	64 (0.8)	59 (0.8)	55 (3.2)	64 (0.7)	60 (0.8)	55 (1.7)
New Jersey	61 (1.0)	57 (1.3)	53 (1.2)	60 (1.1)	58 (1.0)	58 (2.4)
New Mexico	57 (0.8)	51 (0.7)	45 (1.3)	56 (0.6)	52 (0.8)	48 (1.0)
New York	61 (1.2)	55 (1.3)	45 (2.4)	59 (1.5)	55 (1.4)	55 (2.4)
North Carolina	56 (0.9)	49 (0.9)	49 (2.2)	55 (1.0)	51 (0.8)	51 (1.7)
North Dakota	66 (0.7)	64 (0.8)	62 (2.9)	65 (0.9)	64 (0.5)	65 (2.0)
Ohio	61 (1.0)	57 (1.1)	51 (2.7)	60 (1.6)	56 (1.1)	56 (1.5)
Oklahoma	61 (1.0)	55 (0.7)	51 (1.3)	58 (1.0)	57 (0.8)	52 (1.7)
Pennsylvania	61 (1.0)	55 (0.9)	52 (2.4)	60 (1.1)	57 (1.0)	53 (3.0)
Rhode Island	60 (0.5)	51 (0.6)	48 (2.4)	57 (0.6)	55 (0.6)	52 (1.4)
South Carolina	58 (0.9)	51 (0.9)	47 (1.7)	55 (0.8)	53 (1.0)	52 (2.6)
Tennessee	57 (1.0)	51 (0.8)	47 (2.8)	55 (1.0)	52 (0.7)	50 (3.1)
Texas	59 (1.0)	53 (1.1)	47 (2.0)	58 (1.1)	55 (1.0)	51 (1.7)
Utah	62 (0.6)	58 (0.9)	54 (2.0)	62 (0.8)	59 (0.8)	57 (1.5)
Virginia	62 (0.9)	54 (1.1)	49 (1.5)	60 (1.2)	55 (0.9)	50 (1.7)
West Virginia	56 (0.9)	50 (0.9)	46 (1.3)	56 (0.9)	51 (0.8)	48 (1.3)
Wisconsin	64 (1.0)	60 (1.0)	55 (1.9)	64 (1.2)	62 (0.7)	57 (2.2)
Wyoming	63 (0.7)	59 (0.7)	53 (1.1)	64 (0.8)	58 (0.7)	58 (1.5)
TERRITORIES						
Guam	52 (0.9)	41 (0.6)	39 (1.1)	49 (0.7)	45 (0.9)	35 (0.9)
Virgin Islands						

TABLE 3.14 contains the results for a similar analysis for the twelfth graders, which looked at the relationship between their reports of high-school mathematics course-taking and performance by question type. In general, the data reveal a relationship between course-taking and performance for each question type -- the more advanced the course, the higher the students' performance. The data also illustrate the relative difficulty of the three question types. For each level of course work, students' performance improved significantly from their performance on the extended-response questions to their performance on the regular constructed-response questions to their performance on the multiple-choice questions. However, note that even the high-school seniors who reported having studied calculus had difficulty with the extended-response questions. Certainly, these students as well as their classmates in precalculus classes had the mathematical background to have studied the knowledge and procedures required to solve the problems included in the assessment.

TABLE 3.14 Average Percentages of Successful Responses for Different Question Types by Courses Taken in Mathematics

Grade 12

GRADE 12	Not Studied	Prealgebra	First-Year Algebra	Second-Year Algebra	Precalculus	Calculus
Satisfactory or Better on Extended-Response	4(1.1)	3(0.8)	7(0.7)	8(0.6)	13(1.3)	23(1.9)
Correct on Regular Constructed-Response	22(1.3)	26(1.2)	33(1.0)	44(0.7)	54(0.9)	61(1.7)
Correct on Multiple-Choice	38(1.1)	44 (0.9)	50(0.8)	60(0.5)	69(0.7)	74(1.4)

The standard errors of the estimated percentages appear in parentheses. It can be said with about 95 percent certainty for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details).

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment

Summary

As shown in TABLE 3.15, summarizing student performance by type of question, students had great difficulty with the extended-response tasks compared to the regular constructed-response and multiple-choice questions. Also, performance was somewhat lower on the regular constructed-response questions than on the multiple-choice questions.

TABLE 3.15 Summary of Average Percentage Correct* by Type of Questions

	Grade 4			Grade 8			Grade 12		
	Extended-Response	Regular Constructed-Response	Multiple-Choice	Extended-Response	Regular Constructed-Response	Multiple-Choice	Extended-Response	Regular Constructed-Response	Multiple-Choice
Nation	16 (0.6)	42 (0.5)	50 (0.4)	8 (0.5)	53 (0.5)	56 (0.4)	9 (0.4)	40 (0.5)	56 (0.4)
Northeast	21 (1.7)	45 (1.4)	52 (1.1)	10 (1.4)	54 (1.5)	57 (1.3)	10 (0.8)	42 (0.9)	58 (0.6)
Southeast	13 (1.3)	36 (1.2)	46 (0.8)	6 (0.6)	49 (0.8)	53 (0.7)	7 (0.6)	36 (0.8)	54 (0.6)
Central	17 (1.0)	45 (1.0)	51 (1.0)	9 (0.8)	57 (1.1)	59 (0.8)	10 (0.9)	43 (1.2)	58 (0.9)
West	15 (1.1)	42 (1.0)	49 (0.8)	8 (0.9)	53 (1.0)	56 (1.0)	8 (0.7)	40 (1.0)	56 (0.8)
White	20 (0.8)	47 (0.6)	53 (0.5)	10 (0.6)	59 (0.6)	60 (0.5)	10 (0.5)	44 (0.6)	59 (0.4)
Black	5 (0.7)	24 (0.8)	38 (0.6)	2 (0.3)	36 (0.9)	42 (0.6)	4 (0.7)	26 (0.9)	46 (0.9)
Hispanic	7 (1.0)	31 (0.7)	42 (0.7)	3 (0.5)	42 (0.7)	46 (0.7)	4 (0.6)	32 (0.9)	49 (1.0)
Male	16 (0.8)	43 (0.5)	50 (0.5)	7 (0.5)	53 (0.7)	56 (0.6)	8 (0.5)	41 (0.7)	58 (0.5)
Female	17 (0.8)	41 (0.7)	48 (0.5)	10 (0.7)	54 (0.5)	56 (0.5)	9 (0.5)	40 (0.6)	55 (0.5)
Advantaged Urban	26 (2.4)	54 (1.3)	59 (1.4)	16 (2.1)	64 (1.8)	65 (1.9)	13 (1.1)	49 (1.8)	65 (1.1)
Disadvantaged Urban	5 (1.0)	26 (1.4)	38 (1.2)	3 (0.8)	37 (1.5)	43 (1.2)	5 (0.7)	30 (1.1)	48 (1.2)
Extreme Rural	14 (1.9)	40 (2.6)	48 (1.7)	6 (1.0)	53 (2.8)	56 (2.3)	7 (0.8)	37 (1.5)	53 (1.2)
Other	17 (0.7)	42 (0.6)	50 (0.5)	8 (0.5)	54 (0.6)	56 (0.5)	9 (0.5)	41 (0.6)	57 (0.4)
Public	16 (0.7)	41 (0.6)	49 (0.5)	8 (0.5)	52 (0.5)	56 (0.5)	8 (0.4)	39 (0.6)	56 (0.5)
Catholic and Other Private	19 (1.1)	47 (0.9)	53 (0.8)	11 (1.0)	60 (1.3)	62 (1.1)	13 (1.0)	48 (1.5)	63 (1.2)

*Data for Extended-Response questions are for the average percentages of satisfactory or better responses.

The standard errors of the estimated percentages appear in parentheses. It can be said with about 95 percent certainty for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details).

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment

In general, the absolute differences in average performance among subgroups were larger for the regular constructed-response and multiple-choice questions than for the extended-response questions. For example, the difference in average performance favoring White twelfth graders compared to Black twelfth graders was 18 percentage points for regular constructed-response questions, 13 percentage points for multiple-choice questions, and 6 percentage points for extended-response questions. However, the relative difficulty of the various question types detracts from the ability to make such comparisons meaningful. The extended-response questions were so difficult that large differences among subgroups could not be found. That is, because no subgroup performed very well, this floor effect placed a low boundary on the differences that could possibly occur. Also, in comparing results for multiple-choice questions as opposed to constructed-response formats, the fact that students can guess on multiple-choice questions becomes a consideration. At present, there is no good way to take these various factors into account in making comparisons across question types. Nevertheless, differences among some subgroups were consistent across the three types of questions. Even on the extended-response questions, White students had better performance, on average, than did Black and Hispanic students. Also, performance gaps were observed between students attending schools in advantaged urban communities and those in disadvantaged urban communities as well as between those attending private schools and those attending public schools.

Research using the data from grade 8 indicates that the difficulty of the extended-response questions also affects their contribution to summary performance scales developed through item response theory (IRT) methods. For the more proficient students, these types of tasks provide more information than either multiple-choice or regular constructed-response questions. For the less proficient students, however, the regular constructed-response questions provided the most information. Still, overall, the extended-response questions provided the most information of the three types of questions. As students spend more classroom time in performance-based problem-solving situations, performance should improve on more involved assessment tasks.

Instructional emphasis on reasoning as well as students having taken advanced course work is related to higher performance regardless of the type of assessment question. In closing, however, it can be noted that even for those twelfth-grade students who reported having taken calculus, the average percentage of satisfactory or better responses provided for the extended-response questions was just 23 percent.

PROCEDURAL APPENDIX

Overview of Procedures Used in NAEP's 1992 Mathematics Assessment

The Assessment Framework

As described earlier in the report, the framework underlying NAEP's 1992 mathematics assessment was initially developed for the 1990 assessment and subsequently approved for use in both assessments by the National Assessment Governing Board. It was developed through a consensus process managed by the Council of Chief State School Officers, and the items were developed through a similarly broad-based process managed by Educational Testing Service. The development of the mathematics assessments, including the Trial State Assessment Program, benefited from the involvement of hundreds of representatives from State Education Agencies who attended numerous NETWORK meetings; served on committees; reviewed the framework, objectives, and questions; and in general, provided important suggestions on all aspects of the program.

The mathematics assessment framework is a five-by-three matrix specifying five content areas -- Numbers and Operations; Measurement; Geometry; Data Analysis, Statistics, and Probability; and Algebra and Functions, plus three process or ability areas. These include Conceptual Understanding, Procedural Knowledge, and Problem Solving.[16] TABLES A.1 and A.2 show the approximate percentage distribution of questions for the entire assessment by content area, mathematical ability, and grade.

[16] *Mathematics Objectives, 1990 Assessment* (Princeton, NJ: National Assessment of Educational Progress, Educational Testing Service, 1988).

TABLE A.1 Target and Actual Percentage Distribution of Questions by Grade and Content Area

Content Area	Grade 4		Grade 8		Grade 12	
	Target	Actual	Target	Actual	Target	Actual
Numbers and Operations	45	41	30	32	25	25
Measurement	20	19	15	18	15	16
Geometry	15	17	20	20	20	18
Data Analysis, Statistics, and Probability	10	13	15	15	15	16
Algebra and Functions	10	10	20	16	20	25

Actual percentages are based on the regular constructed-response and multiple-choice questions, and do not exclude the extended-response questions (please see the TABLE 1 in the Introduction to this report).

TABLE A.2 Target and Actual Percentage Distribution of Questions by Grade and Mathematical Ability

Mathematical Ability	Grade 4		Grade 8		Grade 12	
	Target	Actual	Target	Actual	Target	Actual
Conceptual Understanding	40	40	40	37	40	39
Procedural Knowledge	30	20	30	24	30	29
Problem Solving	30	40	30	39	30	32

Actual percentages are based on the classifications agreed upon by NAEP's 1992 Item Development Committee. It is recognized that making discrete classifications is difficult for these categories and that independent efforts to classify NAEP questions have led to different results.[17] The *Mathematics Framework* for the 1994 assessment is based on an integrated approach, whereby these mathematical abilities are considered within the context of reasoning, connections, and communications, so that questions can include aspects of both mathematical abilities and mathematical power.[18]

The Assessment Design

Each student received a booklet containing a set of general background questions, a set of subject-specific background questions, three 15-minute segments, or blocks, of cognitive items, and a set of questions about his or her motivation and familiarity with the assessment material. The same booklets were used in both the national and Trial State Assessments. At each grade level, the mathematics assessment included 16 different blocks of multiple-choice and constructed-response content questions. Students received different blocks of

[17] *Assessing Student Achievement in the States.* The First Report of the National Academy of Education Panel on the Evaluation of the NAEP Trial State Assessment: 1990 Trial State Assessment (Stanford, CA: National Academy of Education, 1992).

[18] *1994 National Assessment of Educational Progress: Mathematics Framework* (Washington, DC: National Assessment Governing Board, U.S. Government Printing Office, 1993).

cognitive items in their booklets according to a careful plan. The 1992 assessment was based on an adaptation of matrix sampling called balanced incomplete block (BIB) spiraling -- a design that enables broad coverage of mathematics content while minimizing the burden for any one student. The balanced incomplete block part of the design assigns blocks of items to booklets and each pair of blocks appears together in at least one booklet. The spiraling part of the method cycles the booklets for administration, so that typically only a few students in any assessment session receive the same booklet.

Thirteen of the 16 blocks were assembled in accordance with this design, whereby the 13 blocks were presented in 26 booklets. Each block appeared in exactly six booklets, and each block appeared with every other block in at least one booklet. Students at grades 4 and 8 were given calculators to use with three of the 13 blocks and were trained in their use prior to the assessment. Students at grade 12 were given calculators to use with four of the 13 blocks. At the fourth grade, students were provided with four-function calculators and at grades 8 and 12, they were provided with scientific calculators. For another block, fourth-grade students were provided with a ruler, and eighth- and twelfth-grade students with a protractor/ruler. For still another block, at all three grades, students were given geometric shapes (manipulatives) to provide a concrete basis for determining their answers. For the national assessment, the three remaining blocks at each grade used a paced-audiotape format to measure students' estimation skills and to move students through some experimental materials.

As part of the 1992 mathematics assessments, including the Trial State Assessment Program, questionnaires were given to the mathematics teachers of the fourth- and eighth-grade students participating in the assessment and to the principal or other administrator in each participating school. An expert panel developed guidelines for the school and teacher questionnaires focusing on five educational areas: instructional content, instructional practices and experiences, teacher characteristics, school conditions and contexts, and conditions beyond school (i.e., home support, out-of-school activities, and attitudes).[19]

Because the sampling for the teacher questionnaires was based on participating students, the responses to the mathematics teacher questionnaire do not necessarily represent all fourth- or eighth-grade mathematics teachers in the nation, or in a state or territory. Rather, they represent teachers of the representative sample of students assessed. It is important to note that in this report, as in all NAEP reports, the student is always the unit of analysis, even when information from the teacher or school questionnaire is being reported. Using the student as the unit of analysis makes it possible to describe the instruction received by representative samples of students. Although this

[19]National Assessment of Educational Progress, *1992 Policy Information Framework* (Princeton, NJ: National Assessment of Educational Progress, Educational Testing Service, 1992).

approach may provide a different perspective from that obtained by simply collecting information from teachers or schools, it is consistent with NAEP's goal of providing information about the educational context and performance of students.

National Sampling

Sampling and data collection activities for 1992 NAEP assessments were conducted by Westat, Inc. In 1992, the assessment was conducted from January through March, with some make-up sessions in early April.

As with all NAEP national assessments, the results for the national samples were based on a stratified, three-stage sampling plan. The first stage included defining geographic primary sampling units (PSUs), which are typically groups of contiguous counties, but sometimes a single county; classifying the PSUs into strata defined by region and community type; and randomly selecting PSUs. For each grade, the second stage included listing, classifying, and randomly selecting schools, both public and private, within each PSU selected at the first stage. The third stage involved randomly selecting students within a school for participation. Some students who were selected (about 7 to 8 percent) were excluded because of limited English proficiency or severe disability.

TABLE A.3 presents the student and school sample sizes and the cooperation and response rates for the national assessment.

TABLE A.3 1992 Student and School Sample Sizes

	Number of Participating Schools	Percent of Schools Participating	Number of Students	Percent of Student Completion
Grade				
4	527	86	8,738	93
8	406	84	9,432	89
12	304	81	8,499	81
Total	1,237		26,669	

Although sampled schools that refused to participate were occasionally replaced, school cooperation rates were computed based on the schools originally selected for participation in the assessments. The rates, which are based on schools sampled for all subjects assessed in 1992 (reading, writing, and mathematics) are also the best estimates for the mathematics assessment. The

student completion rates represent the percentage of students assessed of those invited to be assessed in mathematics, including those assessed in follow-up sessions, when necessary. The BIB-spiraled portion of the assessment (13 blocks, 26 booklets) was administered to 7,176 students at grade 4; 7,663 students at grade 8; and 6,973 students at grade 12. The remaining students participated in the estimation study. Of the participating schools, 944 were public schools, and 638 were Catholic and other private schools.

Many of the results presented in this report are for individual questions, rather than summarized across all questions. In particular, the results are for problem-solving tasks contained in the BIB-spiral portion of the assessment. In accordance with this design, each block -- and therefore each question -- was administered to a nationally representative sample of approximately 1,500 to 1,700 students at each of the three grades assessed.

Trial State Assessment Sampling

For the 44 jurisdictions participating in the 1992 Trial State Assessment Program, the basic design for each grade was to select a sample of 100 public schools from each state, with a sample of 30 students drawn from each school. In the eighth grade, up to three sessions (90 students) were selected from large schools to better represent this school type. For states with small numbers of schools, and no or very few small schools, all schools were included in the sample with certainty. In the fourth grade, all the eligible fourth-grade schools in the District of Columbia, Delaware, and Guam were taken into the sample with certainty. In the eighth grade, all the eligible schools were taken from the District of Columbia, Delaware, Hawaii, Rhode Island, Guam, and the Virgin Islands.

In states where a sample of schools was drawn, schools were stratified by urbanicity, minority strata (which varied by state and urbanicity level), and median income. Special procedures were used for small schools and for identifying and including new schools in the sampling frame for each jurisdiction. To minimize the potential for nonresponse bias, substitutes for nonparticipating schools were selected on a one-by-one basis to be similar to the original school in terms of urbanicity, percent Black enrollment, percent Hispanic enrollment, median household income, and total fourth- or eighth-grade enrollment. Furthermore, the substitute school was selected from the same district whenever possible.

In Guam and the Virgin Islands, all grade eligible students were targeted for inclusion in the assessment.[20] In the remaining jurisdictions, a systematic equal probability sample of the desired number of students (usually 30, but

[20] In the Virgin Islands half the fourth graders were assigned to the mathematics assessment and half to reading. In Guam, students participated in both assessments.

sometimes more) was drawn from each school, typically yielding a sample size in excess of 2,500 students at each grade for each participating state and territory. Representative samples of approximately 600 to 700 students at each grade in each participating state and territory responded to each question or task. The state assessments were conducted during February.

Participation Rates for States and Territories

Summary information about school and student participation rates for each state (including the District of Columbia) and territory is contained in TABLE A.4, which also contains comparable information for the national and regional subsamples used in this report as a basis for comparison to states and territories. More specifically, these results are based only on students attending public schools (not private schools). The guidelines for receiving notations about participation are presented below. Consistent with NCES statistical standards,[21] weighted data have been used to calculate all participation rates. A discussion of the variation in participation rates is found in the *Technical Report of the 1992 Trial State Assessment in Mathematics.*

Since 1989, state representatives, the National Assessment Governing Board (NAGB), several committees of external advisors to the National Assessment of Educational Progress (NAEP), and the National Center for Education Statistics (NCES) have engaged in numerous discussions about the procedures for reporting the NAEP Trial State Assessment results. As part of these discussions, it was recognized that sample participation rates across the states and territories have to be uniformly high to permit fair and valid comparisons. Unless the overall participation rate is high for a state or territory, there is a risk that the assessment results for that jurisdiction are subject to appreciable nonresponse bias. Moreover, even if the overall participation rate is high, there may be significant nonresponse bias if the nonparticipation that does occur is heavily concentrated among certain classes of schools or students. Therefore, NCES established four guidelines for school and student participation in the 1990 Trial State Assessment Program.

For the 1992 Trial State Assessment, NCES decided to continue to use those four guidelines, two relating to school participation and two relating to student participation. The guidelines are based on the standards for sample surveys that are set forth in the *NCES Statistical Standards (1992).* Three of the guidelines for the 1992 program are identical to those used in 1990, while one guideline for school participation has been modified.

---

[21] *NCES Statistical Standards,* NCES 92-021 (Washington DC: National Center for Education Statistics, U.S. Department of Education, 1992).

The following notations concerning school and student participation rates in the Trial State Assessment Program were established to address four significant ways in which nonresponse bias could be introduced into the jurisdiction sample estimates. The four conditions that will result in a state or territory receiving a notation in the 1992 reports are presented below. Note that in order to receive no notations, a state or territory must satisfy all the guidelines at both grade 4 and grade 8.

A jurisdiction will receive a notation if:

1. **Both the state's weighted participation rate for the initial sample of schools was below 85 percent AND the weighted school participation rate after substitution was below 90 percent; OR the weighted school participation rate of the initial sample of schools was below 70 percent (regardless of the participation rate after substitution).**

Discussion: For states or territories that did not use substitute schools, the participation rates are based on participating schools from the original sample. In these situations, the NCES standards specify weighted school participation rates of at least 85 percent to guard against potential bias due to school nonresponse. Thus, the first part of the notation that refers to the weighted school participation rate for the initial sample of schools is in direct accordance with NCES standards.

To help ensure adequate sample representation for each jurisdiction participating in the 1992 Trial State Assessment Program, NAEP provided substitutes for nonparticipating schools. When possible, a substitute school was provided for each initially selected school that declined participation before November 15, 1991. For states or territories that used substitute schools, the assessment results will be based on the student data from all participating schools from both the original sample and the list of substitutes (unless both an initial school and its substitute eventually participated, in which case only the data from the initial school will be used).

The NCES standards do not explicitly address the use of substitute schools to replace initially selected schools that decide not to participate in the assessment. However, considerable technical consideration was given to this issue. Even though the characteristics of the substitute schools were matched as closely as possible to the characteristics of the initially selected schools, substitution does not entirely eliminate bias due to the nonparticipation of initially selected schools. Thus, for the weighted school participation rates including substitute schools, the guideline was set at 90 percent.

Finally, if the jurisdiction's school participation rate for the initial sample of schools is below 70 percent, even if the rate after substitution exceeds 90

percent, there is a substantial possibility that, in aggregate, the substitute schools are not sufficiently similar to the schools that they replaced to assure that there is negligible bias in the assessment results. The last part of the notation takes this into consideration.

A jurisdiction will receive a notation if:

2. The nonparticipating schools included a class of schools with similar characteristics, which together accounted for more than five percent of the state's total fourth- or eighth-grade weighted sample of public schools. The classes of schools from each of which a state needed minimum school participation levels were determined by urbanicity, minority enrollment, and median household income of the area in which the school is located.

Discussion: The NCES standards specify that attention should be given to the representativeness of the sample coverage. Thus, if some important segment of the jurisdiction's population is not adequately represented, it is of concern, regardless of the overall participation rate.

This notation addresses the fact that, if nonparticipating schools are concentrated within a particular class of schools, the potential for substantial bias remains, even if the overall level of school participation appears to be satisfactory. Nonresponse adjustment cells have been formed within each jurisdiction, and the schools within each cell are similar with respect to minority enrollment, urbanicity, and/or median household income, as appropriate for each jurisdiction.

If more than five percent (weighted) of the sampled schools (after substitution) are nonparticipants from a single adjustment cell, then the potential for nonresponse bias is too great. This guideline is based on the NCES standard for stratum-specific school nonresponse rates.

A jurisdiction will receive a notation if:

3. The weighted student response rate within participating schools was below 85 percent.

Discussion: This guideline follows the NCES standard of 85 percent for overall student participation rates. The weighted student participation rate is based on all eligible students from initially selected or substitute schools who participated in the assessment in either an initial session or a make-up session. If the rate falls below 85 percent, then the potential for bias due to students' nonresponse is too great.

A jurisdiction will receive a notation if:

4. **The nonresponding students within participating schools included a class of students with similar characteristics, who together comprised more than five percent of the state's weighted assessable student sample. Student groups from which a state needed minimum levels of participation were determined by age of student and type of assessment session (unmonitored or monitored), as well as school urbanicity, minority enrollment, and median household income of the area in which the school is located.**

Discussion: This notation addresses the fact that if nonparticipating students are concentrated within a particular class of students, the potential for substantial bias remains, even if the overall student participation level appears to be satisfactory. Student nonresponse adjustment cells have been formed using the school-level nonresponse adjustment cells, together with the student's age and the nature of the assessment session (unmonitored or monitored). If more than five percent (weighted) of the invited students who do not participate in the assessment are from a single adjustment cell, then the potential for nonresponse bias is too great. This guideline is based on the NCES standard for stratum-specific student nonresponse rates.

Although the first and third guidelines about school and student participation rates were considered most salient in summarizing overall participation rates as presented in TABLE A.5, it should be noted that several participating entities also failed to meet the conditions for participation across classes of schools with similar characteristics specified under guideline 2. Those receiving notations for guideline 2 included Delaware, Maine, Nebraska, New Hampshire, New Jersey, New York, and Guam at grade 4. Maine, New Jersey, and New York did not satisfy guideline 2 at grade 8. All participants met or exceeded guideline 4 about minimum participation rates for classes of students with similar characteristics.

The results of further study of participation rates for entities that failed to meet the guidelines are presented in the *Technical Report of the 1992 Trial State Assessment in Mathematics*. Evidence of significant nonresponse bias was not detected. However, the participation rate data are presented so that readers of the report can accurately assess the quality of the data being presented.

TABLE A.4 | **Summary of School and Student Participation**

PUBLIC SCHOOLS	Grade 4 - 1992					
	Weighted Percentage School Participation Before Substitution	Weighted Percentage School Participation After Substitution	Notation Number 1	Weighted Percentage Student Participation After Make-ups	Notation Number 3	Weighted Overall Rate
NATION	86	86		94		81
Northeast	82	82		94		78
Southeast	94	94		93		88
Central	92	92		94		87
West	79	79		94		75
STATES						
Alabama	75	97		95		93
Arizona	100	100		95		95
Arkansas	90	99		96		95
California	91	97		94		91
Colorado	100	100		95		95
Connecticut	99	99		96		95
Delaware	92	92		95		87
Dist. Columbia	99	99		93		92
Florida	100	100		95		95
Georgia	100	100		95		95
Hawaii	100	100		95		95
Idaho	84	97		97		94
Indiana	76	91		96		87
Iowa	100	100		96		96
Kentucky	93	96		96		92
Louisiana	100	100		95		95
Maine	57	71	***	95		68
Maryland	99	99		96		95
Massachusetts	87	97		95		92
Michigan	83	90		94		84
Minnesota	82	94		95		89
Mississippi	98	100		97		97
Missouri	89	97		96		93
Nebraska	80	87	***	96		83
New Hampshire	69	80	***	96		77
New Jersey	76	82	***	96		79
New Mexico	75	90		95		86
New York	78	83	***	96		80
North Carolina	95	99		95		94
North Dakota	73	90		96		87
Ohio	79	91		95		87
Oklahoma	86	98		84	***	83
Pennsylvania	84	95		96		91
Rhode Island	83	96		95		91
South Carolina	98	99		97		96
Tennessee	92	93		96		89
Texas	93	98		96		94
Utah	99	99		96		95
Virginia	99	99		95		94
West Virginia	100	100		96		96
Wisconsin	100	100		96		96
Wyoming	97	97		96		93
TERRITORY						
Guam	94	94		95		89

See explanations of the notations and guidelines about sample representativeness. Weighted percentages for the nation and region are based on schools sampled for all subject areas assessed in 1992 (mathematics, reading, and writing). However, based on the national sampling design, the rates shown also are the best estimates for the mathematics assessment. **Notation Number 1** = Both the state's weighted participation rate for the initial sample of schools was below 85% AND the weighted school participation rate after substitution was below 90%; OR the weighted school participation rate of the initial sample of schools was below 70% (regardless of the participation rate after substitution.) **Notation number 3** = The weighted student response rate within participating schools was below 85 percent.

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment.

TABLE A.4 | **Summary of School and Student Participation (continued)**

PUBLIC SCHOOLS	Grade 8 - 1992					
	Weighted Percentage School Participation Before Substitution	Weighted Percentage School Participation After Substitution	Notation Number 1	Weighted Percentage Student Participation After Make-ups	Notation Number 3	Weighted Overall Rate
NATION	88	89		89		79
Northeast	92	92		89		82
Southeast	94	94		90		85
Central	86	87		89		78
West	82	84		88		74
STATES						
Alabama	66	92	***	95		88
Arizona	99	99		93		92
Arkansas	89	97		94		91
California	93	98		92		90
Colorado	100	100		93		93
Connecticut	99	99		94		93
Delaware	100	100		92		92
Dist. Columbia	100	100		85		85
Florida	100	100		91		91
Georgia	99	99		93		92
Hawaii	100	100		90		90
Idaho	85	91		95		86
Indiana	79	94		94		88
Iowa	99	99		95		94
Kentucky	96	98		96		94
Louisiana	100	100		92		92
Maine	62	84	***	93		78
Maryland	89	91		92		84
Massachusetts	83	95		94		89
Michigan	78	94		94		88
Minnesota	81	92		94		87
Mississippi	99	100		95		95
Missouri	92	99		95		94
Nebraska	75	85	***	96		81
New Hampshire	80	92		94		86
New Jersey	69	78	***	94		73
New Mexico	77	94		93		87
New York	81	83	***	92		77
North Carolina	94	98		94		92
North Dakota	78	97		96		93
Ohio	77	90		93		83
Oklahoma	82	98		80	***	79
Pennsylvania	81	94		94		89
Rhode Island	85	100		93		92
South Carolina	94	97		94		91
Tennessee	87	91		94		86
Texas	95	99		94		93
Utah	100	100		94		94
Virginia	97	97		94		92
West Virginia	100	100		94		94
Wisconsin	100	100		94		94
Wyoming	99	99		95		94
TERRITORIES						
Guam	100	100		90		90
Virgin Islands	100	100		92		92

Excluded Students

It is NAEP's intent to assess all selected students. Therefore, all selected students who are capable of participating in the assessment should be assessed. However, some students sampled for participation in NAEP are excluded from the sample according to carefully defined criteria. Specifically, some of the students identified as having Limited English Proficiency (LEP) or having an Individualized Education Plan (IEP) may be incapable of participating meaningfully in the assessment. These students are identified as follows:

LEP students may be excluded if:

- The student is a native speaker of a language other than English; AND
- He or she has been enrolled in an English-speaking school for less than two years; AND
- The student is judged to be incapable of taking part in the assessment.

IEP students may be excluded if:

- The student is mainstreamed less than 50 percent of the time in academic subjects and is judged to be incapable of taking part in the assessment, OR
- The IEP team has determined that the student is incapable of taking part meaningfully in the assessment.

When there is doubt, the student is included in the assessment.

For each student excluded from the assessment, including those in the 1992 Trial State Assessment Programs, school personnel complete a questionnaire about the characteristics of that student and the reason for exclusion. Approximately 7 to 8 percent of the students nationally were excluded from the assessment. Across the participating states and territories, the percentages ranged from 2 to 12 percent at grade 4 and from 2 to 10 percent at grade 8.

Data Collection

As with all NAEP assessments, data collection for the 1992 assessment was conducted by a trained field staff. For the national assessment, this was accomplished by Westat staff. However, in keeping with the legislative requirements of the Trial State Assessment Program, the state mathematics assessments involving approximately 111,000 fourth graders and 109,000 eighth graders in about 9,000 schools were conducted by personnel from each of the participating states. NAEP's responsibilities included selecting the sample of schools and students for each participating state, developing the administration procedures and manuals, training the personnel who would conduct the assessments, and conducting an extensive quality assurance program.

Each participating state and territory was asked to appoint a State Coordinator to be the liaison between NAEP and participating schools. The State Coordinator was asked to gain cooperation of the selected schools, assist in scheduling, provide information necessary for sampling, and notify personnel about training. At the local school level, the administrators, usually school or district staff, were responsible for attending training, identifying excluded students, distributing school and teacher questionnaires, notifying sampled students and their teachers, administering the assessment session, completing the necessary paperwork, and preparing the materials for shipment.

Westat staff trained assessment administrators within the states in three and one-half hour sessions that included a videotape and practice exercises to provide uniformity in procedures. Almost 10,000 persons who were to be assessment administrators were trained in about 500 training sessions around the nation.

To provide quality control across states, a randomly selected 50 percent of the state assessment sessions were monitored by approximately 400 quality control monitors, who were also trained Westat staff. The identity of the schools to be monitored was not revealed to state, district, or school personnel until shortly before the assessment was to commence. The analysis of the results for the unmonitored schools as compared to the monitored schools yielded no systematic differences that would suggest different procedures were used. See the *Technical Report of the 1992 Trial State Assessment in Mathematics* for details and results of this analysis.

Scoring

Materials from the 1992 assessment, including the Trial State Assessment Program, were shipped to National Computer Systems in Iowa City for processing. Receipt and quality control were managed through a sophisticated bar-coding and tracking system. After all appropriate materials were received from a school, they were forwarded to the professional scoring area, where the responses to the open-ended items were evaluated by trained staff using guidelines prepared by NAEP. Each open-ended question had a unique scoring guide that defined the criteria to be used in evaluating students' responses. Of the regular constructed-response items, most were scored right/wrong, but some included several different categories of correct and incorrect responses. The extended constructed-response questions were evaluated on a scale of 1 to 5, permitting degrees of partial credit to be given (see Figure 2.1).

For the national mathematics assessment and the Trial State Assessment Program approximately 4 million student responses were scored, including a 20 percent reliability sample. The scoring reliability sample sizes per question for each grade were approximately 600 for the nation and 6,700 for the states. The overall percentage of agreement between readers for both the national and Trial State Assessment reliability samples at each of the three grades assessed was 94 percent. In general, scoring reliabilities for the regular constructed-response questions (discussed in Chapter One) rarely dropped below 90 percent and often approached 98 to 99 percent exact agreement. However, as shown below in TABLE A.5, maintaining high degrees of scorer reliabilities was more difficult for the extended-response tasks discussed in Chapter Two, because the diverse nature of the longer responses and their general lack of clarity made categorization across the five scoring levels a complex task.

TABLE A.5 Percentages of Exact Agreement for Scoring Reliability Samples for Scoring Reliability Samples for Extended-Response Tasks†

Grade 4 - Extended Tasks	Nation	State	Overall
Pizza Comparison	83	82	82
Graphs of Pockets	77	77	77
Laura's Calculator Correction	90	90	90
Compare Geometric Figures*	74	73	73
Number Patterns (Photo Album)*	89	90	89
Grade 8 - Extended Tasks			
Treena's Budget	82	82	82
Marcy's Dot Pattern	86	81	81
Radio Stations	83	79	80
Probability (Leroy's Coins)*	87	85	85
Geometric Shapes (Hallway)*	85	84	84
Number Patterns (Tiles)*	73	69	70
Grade 12 - Extended Tasks			
Effective Tax Rates			91
Patterns of Squares (ending in 5)			92
Graphing Path of Object			86
Bicycle Trip Graph*			73
Center of Disk*			89
Extend Pattern of Tiles*			74

† Based on no response plus 5 categories as described in Chapter Two.
* Unreleased, secure task.

Data Analysis and IRT Scaling

After the assessment information had been compiled in the database, the data were weighted according to the population structure. The weighting for the national and state samples reflected the probability of selection for each student as a result of the sampling design, adjusted for nonresponse. Through poststratification, the weighting assured that the representation of certain subpopulations corresponded to figures from the U.S. Census and the Current Population Survey.[22]

Analyses were then conducted to determine the percentages of students who gave various responses to each cognitive and background question. In determining the percentages of students who gave the various responses to the NAEP cognitive items, a distinction was made between missing responses at the end of each block (i.e., missing responses subsequent to the last item the student

[22] For additional information about the use of weighting procedures in NAEP, see Eugene G. Johnson, "Considerations and Techniques for the Analysis of NAEP Data" in *Journal of Educational Statistics* (December 1989).

answered) and missing responses prior to the last observed response. Missing responses before the last observed response were considered intentional omissions. Missing responses at the end of the block were considered "not-reached," and treated as if they had not been presented to the student. In calculating percentages for each item, only students classified as having been presented the item were included in the denominator of the statistic.

It is standard practice at ETS to treat all nonrespondents to the last item as if they had not reached the item. For multiple-choice and standard constructed-response items, the use of such a convention most often produces a reasonable pattern of results in that the proportion reaching the last item is not dramatically smaller than the proportion reaching the next-to-last item. However, for the blocks that ended with extended-response tasks, use of the standard ETS convention resulted in an extremely large drop in the proportion of students attempting the final item. A drop of such magnitude seemed somewhat implausible. Therefore, for blocks ending with an extended-response task, students who answered the next-to-last item but did not respond to the extended-response task were classified as having intentionally omitted the last item.

Item response theory (IRT) was used to estimate average scale-score proficiency for the nation, various subgroups of interest within the nation, and for the states and territories. IRT models the probability of answering an item correctly as a mathematical function of proficiency or skill. The main purpose of IRT analysis is to provide a common scale on which performance can be compared across groups, such as those defined by grades, and subgroups, such as those defined by race/ethnicity or gender. Because of the BIB-spiraling design used by NAEP, students do not receive enough questions about a specific topic to provide reliable information about individual performance. Traditional test scores for individual students, even those based on IRT, would lead to misleading estimates of population characteristics, such as subgroup means and percentages of students at or above a certain proficiency level. Instead, NAEP constructs sets of plausible values designed to represent the distribution of proficiency in the population. A plausible value for an individual is not a scale score for that individual but may be regarded as a representative value from the distribution of potential scale scores for all students in the population with similar characteristics and identical patterns of item response. Statistics describing performance on the NAEP proficiency scale are based on these plausible values.

They estimate values that would have been obtained had individual proficiencies been observed -- that is, had each student responded to a sufficient number of cognitive items so that proficiency could be precisely estimated.[23]

For the 1992 assessment, a scale ranging from 0 to 500 was created to report performance for each content area. (Numbers and Operations; Measurement; Geometry; Data Analysis, Statistics, and Probability; Algebra and Functions) and for the estimation skill area. The scales summarize examinee performance across all three question types used in the assessment (multiple-choice, regular constructed-response, and extended-response). In producing the scales, three distinct IRT models were used. Multiple-choice items were scaled using the three-parameter logistic (3PL) model; regular constructed-response questions were scaled using the two-parameter logistic (2PL) model; and the extended-response tasks were scaled using a generalized partial-credit (GPC) model (Muraki, 1992).[24] Each scale was based on the distribution of student performance across all three grades assessed in the 1990 national assessment (grades 4, 8, and 12) and had a mean of 250 and a standard deviation of 50. A composite scale was created as an overall measure of students' mathematics proficiency. The composite scale was a weighted average of the five content-area scales, where the weight for each content area was proportional to the relative importance assigned to the content area in the specifications developed by the Mathematics Objectives Panel as shown previously in TABLE A.1. As described earlier, the NAEP proficiency scales make it possible to examine relationships between students' performance and a variety of background factors measured by NAEP. The fact that a relationship exists between achievement and another variable, however, does not reveal the underlying cause of the relationship, which may be influenced by a number of other variables. Similarly, the assessments do not capture the influence of unmeasured variables. The results are most useful when they are considered in combination with other knowledge about the student population and the educational system, such as trends in instruction, changes in the school-age population, and societal demands and expectations.

---

[23] For theoretical justification of the procedures employed, see Robert J. Mislevy, "Randomization-Based Inferences About Latent Variables from Complex Samples," *Psychometrika*, 56(2), 177-196, 1988).

For computational details, see *Focusing the New Design: NAEP 1988 Technical Report* (Princeton, NJ: Educational Testing Service, National Assessment of Education Progress, 1990) and the *1990 NAEP Technical Report*.

[24] Muraki, E., "A Generalized Partial Credit Model: Application of an EM algorithm", *Applied Psychological Measurement*, 16(2), 159-176, 1992.

Per-Item Information Analyses Using IRT

The per-item information analyses presented in Chapter Three were taken from Mazzeo, Yamamoto, & Kulick (1993) and details on the methods used are contained therein. The following contains a brief synopsis of the methods used for the per-item information analyses.

An often used IRT-based indicator of the measurement precision of an item is the so-called item information function (Lord and Novick, 1968).[25] For the 2PL and 3PL models, the item information function is defined as:

$$I_j(\theta) = \frac{P_j'^2}{P_j Q_j} \qquad (1)$$

where P_j is the conventional logistic IRT model (2PL or 3PL), Q_j is (1-P_j), and P'_j is the first derivative with respect to θ (the proficiency scale). As shown in Donoghue (1992),[26] the information function for item j under the GPC model is given by:

$$I_j(\theta) = D^2 a_j^2 [\sum_{k=0}^{m_j} k^2 P_{jk}(\theta) - (\sum_{k=0}^{m_j} k P_{jk}(\theta))^2] \qquad (2)$$

where P_{jk} is the item category characteristic curve for the k^{th} score category and a_j is a slope parameter for item j. Under the IRT assumptions of local independence, the total information function for any group of n items is the sum of the item information functions:

---

[25] John Mazzeo, Kentaro Yamamoto, and Edward Kulick, "Extended Constructed-Response Items in the 1992 NAEP: Psychometrically Speaking Were They Worth the Price?" Paper presented at the 1993 annual meeting of the National Council on Measurement in Education, Atlanta, GA.

F.M. Lord and M.R. Novick, *Statistical Theories of Mental Test Scores* (Reading, MA: Addison-Wesley Publishing Co., 1968).

[26] John R. Donoghue, *An Empirical Examination of the IRT Information in Polytomously Scored Reading Items.* ETS Research Report (No. 93-12), (Princeton, NJ: Educational Testing Service, 1992).

$$I^{(n)}(\theta) = \sum_{j=1}^{n} I_j(\theta) \qquad (3)$$

In order to obtain per-item information, it is natural to consider dividing (3) by the number of items:

$$I(\theta) = \frac{I^{(n)}(\theta)}{n} \qquad (4)$$

The information function defined in (4) provides a local model-based measure of per-item information at each level of θ. It is also useful to have an overall summary measure of per-item information. One summary measure, referred to by Donoghue as expected information, can be obtained by integrating I(θ) with respect to a distribution of θ (f(θ)):

$$E(I) = \int_{\theta} I(\theta) f(\theta) d\theta. \qquad (5)$$

A discrete approximation to f(θ) was used to calculate expected per-item information. A set of Q equally spaced θ values were selected (denoted as $\{X_q, q=1,2,...,Q\}$ and a corresponding set of estimates of $f(X_q)$ were obtained (denoted as $\{w_q, q=1,2,...,Q\}$). E(I) was then approximated by:

$$E(I) \approx \sum_{q=1}^{Q} w_q I(X_q) \qquad (6)$$

E(I) can be thought of as the average per-item information for a group of examinees with θ distribution given by f(θ). As such, it reflects not only the measurement qualities of the items in question but how well they are targeted to the proficiency distribution for the group whose proficiencies they are designed to measure. In the current study, E(I) was calculated with an approximation to normal proficiency distribution.

Using the grade 8 data from the 1992 NAEP mathematics assessment, a single IRT-calibration was carried out for the full item pool. Three separate estimates of per-item information functions were obtained, using these unidimensional item parameter estimates, one for the set of multiple-choice items, one for the set of regular constructed-response items, and one for the extended-

response items. Using a discrete approximation to $f(\theta)=N(0,1)$, where the w_q were obtained at 41 equally spaced θ values from -4 to 4 (the $\{X_q\}$ values), and using the estimated unidimensional item parameters, separate estimates of the expected per-item information were also obtained for the multiple-choice, short constructed-response, and extended-response item sets.

The purpose of producing a common unidimensional scale is to allow direct comparisons of information for the three item types that are not complicated by possible differences in metrics associated with different scales. However, the results suggest that the different mathematics item types might be measuring slightly different aspects of the achievement domain. Calibrating the collection of items together on a unidimensional scale will produce a sort of composite of these different aspects. However, because of the large differences in numbers of items, the composite scale may be somewhat dominated by the item types with larger numbers, in particular the multiple-choice items. Consequently, if lesser amounts of information are provided by the smaller scales, in particular the extended-response scales, this may be partly an artifact of their measuring a slightly different dimension.

In order to check on the sensitivity of results to unidimensional assumptions, per-item information was also obtained based on the separate item-type calibrations from the unidimensional calibration. Analogous information functions were also obtained using the item parameter estimates obtained by carrying out separate IRT calibrations for the multiple-choice, short constructed-response, and extended-response items. In addition, expected per-item information was calculated using the separate set of item parameters generated for each item type. While the item parameter estimates from these separate IRT calibrations are not strictly on the same scale, they are on scales that have been standardized to have the same mean and standard deviation (i.e., 0 and 1, respectively) for the common population of examinees administered the assessment. Hence, comparison of the results for the separate item type scales with the results for the unidimensional scales provides a heuristic reasonableness check on the results of the unidimensional analysis. No substantive differences were found between these two sets of analyses. Hence, the results of the analyses using the unidimensional parameters (transformed to the 0 to 500 NAEP reporting scale) are reported in Chapter Three.

Linking the Trial State Results to the National Results

Although the assessment booklets used in the Trial State Assessment Program were identical to those used in the national assessment, the various differences between the national and trial state assessments, including those in administration procedures, required that careful and complex equating procedures based on a special design be used to create an appropriate basis for comparison between the national and state results.

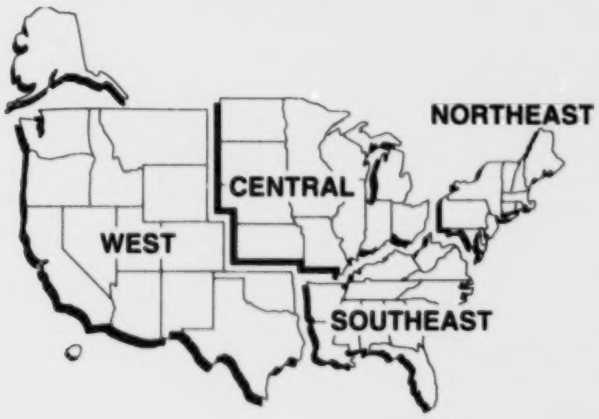
Two separate sets of IRT-based scales (one set based on data from the Trial State Assessment and one set based on national assessment data) were established for the 1992 assessment. The scales from the Trial State Assessment were linked to those from the national assessment through a linking function determined by comparing the results for the aggregate of students assessed in the Trial State Assessment (except those in Guam and the Virgin Islands) with the results for students in the State Aggregate Comparison subsample of the national assessment. This subsample is representative of the population of all grade-eligible public-school students within the aggregate of the 41 participating states and the District of Columbia who were assessed as part of the national assessment.

The linking was accomplished for each subscale by matching the mean and standard deviation of the subscale proficiencies across all students in the Trial State Assessment (excluding Guam and the Virgin Islands) to the corresponding subscale mean and standard deviation across all students in the State Aggregate Comparison subsample.

NAEP Reporting Groups

This report contains results for the nation, participating states, and groups of students within the nation defined by shared characteristics. The definitions for subgroups as defined by region, race/ethnicity, gender, size and type of community, and type of school follow.

Region. The United States has been divided into four regions: Northeast, Southeast, Central, and West. States in each region are shown on the following map.



Race/Ethnicity. Results are presented for students of different racial/ethnic groups based on the students' self-identification of race/ethnicity according to the following mutually exclusive categories: White, Black, Hispanic, Asian/Pacific Islander, and American Indian (including Alaskan Native). Based on statistically determined criteria, at least 62 students in a particular subpopulation must participate in order for the results for that subpopulation to be considered reliable. However, the data for all students, regardless of whether their racial/ethnic group was reported separately, were included in computing the overall national or state level results.

Gender. Results are reported separately for males and females. Gender was reported by the student.

Type of Community. Results are provided for four mutually exclusive community types -- advantaged urban, disadvantaged urban, extreme rural, and other -- as described below. According to information about parents' occupation obtained from the Principal's Questionnaire completed by each sampled school, indices are developed such that for each assessment approximately the 10 percent of the most extreme advantaged urban, disadvantaged urban, and rural schools are classified into the first three categories. The remaining approximately 70 percent of the schools are classified into the "other" category.

Advantaged Urban: Students in this group reside in metropolitan statistical areas and attend schools where a high proportion of the students' parents are in professional or managerial positions.

Disadvantaged Urban: Students in this group reside in metropolitan statistical areas and attend schools where a high proportion of the students' parents are on welfare or are not regularly employed.

Extreme Rural: Students in this group do not reside in metropolitan statistical areas. They attend schools in areas with a population below 10,000 where many of the students' parents are farmers or farm workers.

Other: Students in the "Other" category attend schools in areas other than those defined as advantaged urban, disadvantaged urban, or extreme rural.

Type of School. For the nation, results are presented separately for public-school students and for private-school students, both those attending Catholic schools and other types of private schools combined.

The percentages of students in the national reporting groups are presented in TABLE A.6. Although in this report, state results are not presented separately for subpopulations of students because the question by question sample sizes would be very small, TABLE A.7 contains the characteristics of students by race/ethnicity and type of community. This provides further context in making state-to-nation and state-to-state comparisons.

TABLE A.6 Percentages of Students in Reporting Groups for the Nation

	Grade 4	Grade 8	Grade 12
Northeast	21 (0.9)	22 (0.8)	24 (0.6)
Southeast	24 (0.9)	25 (0.7)	24 (0.6)
Central	27 (0.5)	25 (0.6)	25 (0.6)
West	28 (0.7)	28 (0.7)	27 (0.9)
White	70 (0.2)	70 (0.2)	71 (0.6)
Black	16 (0.1)	16 (0.1)	15 (0.4)
Hispanic	10 (0.2)	10 (0.2)	10 (0.5)
Male	50 (0.6)	51 (0.6)	49 (0.8)
Female	50 (0.6)	49 (0.6)	51 (0.8)
Advantaged Urban	12 (1.8)	10 (1.8)	12 (2.1)
Disadvantaged Urban	9 (1.4)	9 (1.3)	10 (1.4)
Extreme Rural	12 (2.2)	9 (2.6)	12 (1.6)
Other	66 (3.0)	72 (3.1)	66 (3.0)
Public	87 (1.0)	89 (0.9)	87 (1.2)
Catholic and Other Private	13 (1.0)	11 (0.9)	13 (1.2)

The standard erros of the estimated percentages appear in parentheses. It can be said with 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimated for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details). Percentages may not total 100 percent due to rounding error. For the racial/ethnic classifications, small percentages of students reported other categories.

TABLE A.7 | **Characteristics of NAEP Students by Race/Ethnicity and by Type of Community**

PUBLIC SCHOOLS	Grade 4 - 1992								
	Percentage of Students by Race/Ethnicity					Percentage of Students by Type of Community			
	White	Black	Hispanic	Asian / Pacific Islander	American Indian	Advantaged Urban	Disadvantaged Urban	Extreme Rural	Other
NATION	69 (0.4)	17 (0.4)	10 (0.2)	3 (0.3)	2 (0.2)	9 (1.8)	10 (1.5)	13 (2.4)	67 (3.2)
Northeast	71 (2.9)	17 (2.7)	8 (1.2)	2 (0.7)	1 (0.3)	20 (5.5)	16 (5.5)	4 (1.2)	60 (8.0)
Southeast	61 (2.5)	30 (2.6)	6 (1.0)	1 (0.3)	1 (0.3)	5 (3.0)	13 (3.5)	19 (6.9)	63 (7.6)
Central	80 (1.8)	12 (1.7)	6 (0.8)	1 (0.2)	1 (0.3)	5 (2.1)	9 (1.9)	16 (3.4)	70 (4.1)
West	64 (1.7)	10 (1.7)	17 (1.6)	5 (1.0)	2 (0.3)	8 (3.7)	5 (1.3)	13 (4.7)	74 (5.7)
STATES									
Alabama	61 (2.5)	32 (2.3)	4 (0.6)	1 (0.2)	2 (1.0)	11 (3.1)	13 (3.2)	14 (4.0)	62 (5.6)
Arizona	56 (2.1)	4 (0.7)	29 (1.5)	1 (0.2)	10 (1.7)	13 (3.9)	10 (3.0)	8 (3.3)	69 (5.5)
Arkansas	69 (1.5)	21 (1.4)	6 (0.6)	1 (0.2)	3 (0.4)	1 (1.2)	6 (1.5)	25 (4.1)	68 (4.7)
California	45 (2.0)	6 (0.7)	35 (1.7)	11 (1.1)	3 (0.5)	12 (2.5)	23 (3.7)	1 (0.3)	65 (4.6)
Colorado	68 (1.5)	5 (1.0)	22 (1.3)	3 (0.3)	3 (0.3)	18 (3.2)	13 (2.9)	13 (2.7)	57 (5.0)
Connecticut	73 (1.4)	10 (1.1)	13 (1.1)	2 (0.4)	1 (0.2)	19 (4.2)	15 (3.0)	0 (0.0)	66 (5.0)
Delaware	66 (1.1)	23 (0.9)	8 (0.4)	1 (0.2)	2 (0.4)	10 (0.2)	8 (0.2)	24 (0.1)	58 (0.3)
Dist. Columbia	5 (0.4)	82 (0.6)	10 (0.4)	1 (0.2)	2 (0.3)	20 (0.3)	60 (0.4)	0 (0.0)	20 (0.3)
Florida	58 (2.2)	21 (2.0)	17 (1.3)	2 (0.4)	2 (0.3)	18 (4.4)	21 (3.9)	4 (1.3)	57 (4.5)
Georgia	56 (2.2)	35 (2.1)	6 (0.6)	1 (0.2)	1 (0.3)	10 (3.4)	15 (4.6)	12 (3.6)	63 (6.2)
Hawaii	21 (1.6)	4 (0.6)	11 (0.7)	61 (2.1)	2 (0.3)	12 (3.6)	9 (1.8)	5 (1.9)	75 (4.3)
Idaho	84 (1.2)	1 (0.2)	11 (1.0)	1 (0.2)	3 (0.3)	9 (2.6)	1 (0.9)	33 (4.9)	56 (5.5)
Indiana	82 (1.5)	10 (1.3)	5 (0.6)	1 (0.2)	1 (0.3)	8 (2.7)	10 (2.8)	15 (3.3)	68 (4.9)
Iowa	90 (0.9)	2 (0.5)	5 (0.5)	1 (0.3)	2 (0.3)	7 (2.9)	6 (2.5)	41 (3.5)	46 (4.2)
Kentucky	85 (1.6)	9 (1.3)	4 (0.6)	1 (0.2)	2 (0.3)	6 (2.7)	11 (2.7)	24 (4.2)	60 (4.8)
Louisiana	50 (2.0)	43 (2.0)	5 (0.6)	2 (0.7)	1 (0.3)	5 (2.3)	18 (2.5)	11 (2.7)	65 (3.9)
Maine	91 (0.7)	1 (0.1)	5 (0.6)	1 (0.2)	3 (0.5)	2 (1.6)	2 (1.3)	19 (4.7)	77 (4.9)
Maryland	59 (1.7)	30 (1.4)	6 (0.6)	4 (0.5)	2 (0.2)	20 (3.6)	16 (4.0)	5 (2.1)	59 (4.9)
Massachusetts	79 (1.6)	7 (0.8)	8 (0.8)	4 (0.7)	2 (0.2)	16 (3.4)	14 (2.7)	1 (0.9)	68 (4.2)
Michigan	73 (1.8)	13 (1.7)	9 (0.9)	2 (0.3)	3 (0.4)	10 (3.0)	15 (3.7)	10 (3.6)	65 (5.1)
Minnesota	85 (1.3)	3 (0.5)	7 (0.8)	2 (0.4)	2 (0.3)	12 (3.9)	3 (2.2)	29 (3.8)	56 (5.4)
Mississippi	40 (2.0)	52 (2.1)	6 (0.9)	1 (0.2)	1 (0.2)	1 (1.1)	6 (1.9)	11 (2.3)	82 (3.2)
Missouri	77 (1.7)	14 (1.7)	6 (0.5)	1 (0.2)	2 (0.4)	9 (3.0)	11 (2.9)	26 (3.9)	53 (5.3)
Nebraska	84 (1.3)	6 (0.7)	7 (0.9)	1 (0.2)	2 (0.3)	8 (2.7)	6 (1.4)	26 (3.9)	59 (4.8)
New Hampshire	89 (1.2)	1 (0.2)	5 (0.6)	1 (0.2)	3 (0.3)	8 (3.5)	1 (1.3)	4 (1.8)	86 (4.0)
New Jersey	66 (2.2)	14 (1.2)	14 (1.5)	5 (0.8)	1 (0.3)	30 (4.3)	17 (3.3)	1 (1.0)	53 (5.0)
New Mexico	44 (2.4)	4 (0.5)	47 (2.0)	1 (0.3)	4 (1.3)	11 (5.7)	9 (2.9)	4 (2.0)	77 (6.1)
New York	59 (2.2)	13 (1.6)	22 (1.7)	4 (0.8)	2 (0.4)	15 (3.7)	24 (3.7)	2 (1.6)	58 (4.7)
North Carolina	62 (1.7)	29 (1.3)	6 (0.7)	1 (0.2)	3 (0.9)	5 (1.6)	4 (1.9)	19 (4.9)	71 (4.6)
North Dakota	91 (1.0)	0 (0.2)	4 (0.6)	1 (0.2)	4 (0.8)	11 (3.1)	2 (1.4)	43 (3.6)	44 (4.3)
Ohio	79 (1.5)	11 (1.2)	6 (0.5)	1 (0.3)	2 (0.4)	10 (2.6)	18 (2.6)	17 (3.9)	55 (4.8)
Oklahoma	73 (1.5)	9 (1.2)	7 (0.8)	1 (0.2)	10 (0.8)	9 (3.1)	10 (2.6)	21 (3.6)	60 (4.6)
Pennsylvania	77 (1.6)	12 (1.6)	7 (0.8)	2 (0.4)	1 (0.3)	15 (4.9)	17 (3.4)	14 (3.8)	54 (5.6)
Rhode Island	78 (2.1)	6 (1.0)	11 (1.1)	3 (0.4)	2 (0.3)	12 (4.0)	24 (4.9)	0 (0.0)	64 (5.7)
South Carolina	55 (1.7)	37 (1.8)	6 (0.8)	1 (0.2)	1 (0.3)	6 (2.2)	6 (1.5)	13 (3.1)	74 (4.0)
Tennessee	69 (2.1)	23 (1.9)	5 (0.8)	1 (0.4)	1 (0.2)	6 (2.7)	13 (3.6)	10 (2.8)	71 (4.6)
Texas	49 (1.8)	14 (1.8)	34 (2.3)	2 (0.4)	1 (0.2)	10 (3.2)	21 (4.8)	13 (3.3)	56 (6.3)
Utah	86 (1.0)	1 (0.2)	10 (0.8)	2 (0.3)	2 (0.3)	20 (3.6)	3 (1.7)	7 (2.6)	70 (4.4)
Virginia	67 (1.4)	23 (1.3)	5 (0.6)	3 (0.4)	1 (0.3)	13 (3.1)	14 (3.1)	13 (2.7)	59 (4.7)
West Virginia	90 (0.9)	3 (0.4)	5 (0.8)	1 (0.2)	2 (0.2)	2 (1.4)	8 (2.5)	16 (3.6)	75 (4.6)
Wisconsin	81 (1.4)	6 (1.0)	7 (0.7)	2 (0.5)	3 (1.1)	9 (2.6)	7 (2.4)	26 (5.0)	58 (5.3)
Wyoming	82 (1.4)	1 (0.2)	11 (0.9)	1 (0.2)	5 (1.2)	7 (2.1)	4 (1.8)	20 (3.4)	69 (4.5)
TERRITORY									
Guam	12 (0.7)	4 (0.4)	20 (0.8)	62 (1.0)	2 (0.4)	0 (0.0)	0 (0.0)	19 (0.1)	81 (0.1)

The standard errors of the estimated percentages appear in parentheses. It can be said with 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (as described in this Appendix). Percentages may not total 100 percent due to rounding error.

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment.

TABLE A.7 | **Characteristics of NAEP Students by Race/Ethnicity and by Type of Community (continued)**

PUBLIC SCHOOLS	Grade 8 - 1992								
	Percentage of Students by Race/Ethnicity					Percentage of Students by Type of Community			
	White	Black	Hispanic	Asian / Pacific Islander	American Indian	Advantaged Urban	Disadvantaged Urban	Extreme Rural	Other
NATION	69 (0.4)	16 (0.2)	10 (0.3)	2 (0.2)	1 (0.2)	8 (2.2)	9 (1.5)	10 (2.8)	72 (3.5)
Northeast	67 (2.6)	19 (1.5)	10 (1.7)	2 (0.5)	1 (0.3)	12 (6.5)	12 (3.7)	7 (4.8)	69 (8.2)
Southeast	68 (1.8)	27 (1.8)	4 (0.7)	1 (0.3)	1 (0.2)	5 (3.5)	9 (2.5)	16 (7.2)	69 (7.9)
Central	79 (2.0)	13 (1.9)	5 (0.8)	2 (0.5)	1 (0.4)	8 (2.4)	9 (3.0)	9 (6.0)	74 (6.9)
West	63 (1.5)	8 (1.3)	21 (1.7)	5 (0.8)	2 (0.7)	7 (4.0)	9 (3.2)	8 (4.0)	76 (5.3)
STATES									
Alabama	61 (2.3)	32 (2.1)	4 (0.6)	1 (0.2)	2 (0.4)	4 (2.4)	16 (3.5)	15 (3.2)	65 (4.7)
Arizona	60 (2.1)	4 (0.5)	28 (1.6)	2 (0.3)	6 (1.3)	15 (5.3)	14 (3.1)	7 (2.2)	64 (5.8)
Arkansas	72 (1.4)	22 (1.3)	4 (0.4)	1 (0.2)	1 (0.2)	2 (1.4)	5 (1.9)	16 (3.9)	76 (4.4)
California	44 (1.8)	7 (1.1)	36 (1.7)	11 (1.0)	1 (0.2)	8 (3.2)	19 (3.2)	3 (1.9)	71 (5.1)
Colorado	74 (1.2)	4 (0.6)	18 (1.1)	2 (0.3)	2 (0.3)	18 (3.5)	10 (2.3)	13 (2.9)	60 (4.9)
Connecticut	72 (1.6)	12 (1.1)	12 (0.9)	3 (0.4)	0 (0.1)	10 (3.5)	17 (3.3)	0 (0.0)	72 (4.4)
Delaware	65 (0.9)	25 (1.1)	6 (0.6)	2 (0.3)	2 (0.3)	0 (0.0)	0 (0.0)	11 (0.1)	89 (0.1)
Dist. Columbia	3 (0.2)	85 (0.8)	10 (0.7)	1 (0.2)	1 (0.3)	7 (0.3)	67 (0.4)	0 (0.0)	25 (0.4)
Florida	56 (2.1)	23 (2.0)	18 (2.0)	2 (0.3)	1 (0.2)	7 (2.9)	17 (3.5)	6 (2.1)	69 (4.9)
Georgia	59 (2.1)	35 (1.9)	4 (0.5)	2 (0.3)	0 (0.1)	6 (1.9)	10 (2.9)	9 (2.2)	74 (4.0)
Hawaii	17 (0.9)	3 (0.3)	11 (0.7)	66 (1.1)	1 (0.2)	5 (0.1)	16 (0.4)	1 (0.0)	78 (0.4)
Idaho	88 (0.7)	1 (0.2)	8 (0.6)	1 (0.2)	3 (0.4)	4 (2.2)	5 (2.4)	29 (4.3)	62 (5.0)
Indiana	85 (1.3)	8 (1.1)	4 (0.6)	1 (0.2)	1 (0.2)	5 (2.3)	11 (2.4)	13 (2.6)	71 (4.3)
Iowa	92 (0.7)	2 (0.4)	4 (0.4)	1 (0.2)	1 (0.2)	4 (2.3)	3 (1.0)	44 (5.4)	49 (5.7)
Kentucky	87 (1.0)	9 (1.0)	3 (0.4)	1 (0.2)	1 (0.2)	3 (1.1)	12 (3.3)	15 (3.7)	70 (5.1)
Louisiana	54 (1.7)	39 (1.5)	5 (0.5)	2 (0.4)	1 (0.2)	2 (1.6)	19 (3.2)	7 (3.0)	72 (4.3)
Maine	94 (0.5)	0 (0.1)	2 (0.3)	1 (0.2)	3 (0.4)	1 (1.5)	2 (1.6)	19 (4.1)	78 (4.5)
Maryland	60 (1.8)	29 (1.8)	6 (0.6)	3 (0.5)	1 (0.2)	21 (3.8)	13 (3.5)	3 (2.6)	63 (5.6)
Massachusetts	83 (1.1)	5 (1.0)	8 (1.5)	2 (0.4)	1 (0.2)	7 (2.3)	23 (3.5)	1 (1.3)	69 (4.3)
Michigan	73 (1.6)	18 (1.9)	5 (0.8)	1 (0.3)	2 (0.3)	7 (3.0)	19 (3.1)	14 (3.8)	60 (5.2)
Minnesota	91 (1.0)	2 (0.3)	3 (0.5)	2 (0.3)	1 (0.4)	7 (3.6)	0 (0.0)	20 (4.2)	72 (5.2)
Mississippi	49 (1.9)	44 (1.8)	6 (0.6)	0 (0.1)	1 (0.2)	3 (1.8)	6 (2.7)	12 (3.1)	79 (4.6)
Missouri	82 (1.5)	12 (1.4)	3 (0.3)	1 (0.2)	2 (0.3)	7 (2.8)	12 (2.4)	13 (3.6)	68 (4.8)
Nebraska	87 (1.1)	5 (0.9)	6 (0.7)	1 (0.2)	2 (0.4)	0 (0.0)	6 (0.9)	28 (4.3)	66 (4.5)
New Hampshire	91 (1.6)	1 (0.2)	3 (0.3)	1 (0.2)	1 (0.2)	4 (1.6)	0 (0.0)	5 (2.3)	92 (2.8)
New Jersey	61 (2.5)	17 (2.4)	14 (1.5)	6 (0.7)	1 (0.2)	8 (2.8)	24 (3.3)	3 (2.3)	64 (4.7)
New Mexico	44 (1.5)	2 (0.4)	49 (1.4)	1 (0.3)	4 (0.7)	5 (0.2)	6 (2.6)	6 (2.8)	84 (3.8)
New York	61 (2.7)	17 (2.2)	14 (2.0)	4 (0.6)	1 (0.3)	11 (3.3)	16 (5.1)	10 (3.5)	63 (6.7)
North Carolina	68 (1.4)	27 (1.3)	3 (0.3)	1 (0.2)	2 (0.4)	3 (1.0)	5 (2.2)	12 (3.8)	80 (4.3)
North Dakota	93 (0.8)	0 (0.1)	3 (0.3)	1 (0.2)	3 (0.7)	8 (1.8)	0 (0.0)	39 (4.1)	53 (3.9)
Ohio	80 (1.9)	14 (1.7)	4 (0.5)	1 (0.2)	2 (0.3)	6 (2.7)	17 (3.2)	21 (5.5)	56 (6.3)
Oklahoma	75 (1.6)	8 (1.1)	6 (0.6)	2 (0.3)	10 (1.0)	2 (1.8)	5 (2.5)	19 (4.1)	74 (5.1)
Pennsylvania	83 (1.4)	11 (1.6)	3 (0.7)	1 (0.3)	1 (0.3)	4 (2.1)	15 (3.5)	13 (3.7)	68 (5.0)
Rhode Island	81 (0.7)	6 (0.6)	8 (0.4)	3 (0.4)	2 (0.3)	7 (0.1)	12 (0.1)	0 (0.0)	81 (0.1)
South Carolina	58 (1.5)	35 (1.3)	6 (0.6)	1 (0.2)	1 (0.2)	3 (1.7)	6 (2.2)	4 (1.8)	87 (3.3)
Tennessee	75 (2.0)	21 (2.1)	3 (0.3)	0 (0.1)	1 (0.2)	5 (3.3)	7 (2.6)	6 (2.4)	82 (4.0)
Texas	48 (1.9)	12 (1.6)	36 (2.0)	3 (0.4)	1 (0.3)	10 (2.9)	18 (3.9)	6 (2.6)	67 (5.5)
Utah	90 (0.9)	1 (0.2)	7 (0.6)	2 (0.3)	2 (0.2)	13 (2.4)	5 (2.2)	10 (2.4)	72 (3.9)
Virginia	69 (1.9)	22 (1.6)	5 (0.6)	4 (0.5)	1 (0.2)	9 (2.4)	13 (3.0)	14 (4.3)	63 (5.4)
West Virginia	91 (0.9)	4 (0.8)	3 (0.3)	0 (0.1)	2 (0.3)	1 (0.9)	10 (1.9)	13 (3.4)	76 (3.7)
Wisconsin	86 (1.7)	7 (1.2)	4 (0.8)	1 (0.2)	2 (0.6)	11 (5.5)	5 (1.7)	25 (5.4)	59 (6.4)
Wyoming	86 (1.7)	1 (0.2)	9 (0.6)	1 (0.2)	4 (1.6)	0 (0.0)	10 (2.6)	13 (2.9)	76 (3.8)
TERRITORIES									
Guam	5 (0.5)	1 (0.3)	15 (0.9)	76 (1.1)	1 (0.1)	0 (0.0)	0 (0.0)	11 (0.2)	89 (0.2)
Virgin Islands	1 (0.4)	77 (1.1)	21 (0.9)	0 (0.1)	0 (0.2)	0 (0.0)	0 (0.0)	27 (0.2)	73 (0.2)

Estimating Variability

Because the statistics presented in this report are estimates of group and subgroup performance based on samples of students, rather than the values that could be calculated if every student in the nation answered every question, it is important to have measures of the degree of uncertainty of the estimates. Two components of uncertainty are accounted for in the variability of statistics based on proficiency: the uncertainty due to sampling only a relatively small number of students and the uncertainty due to sampling only a relatively small number of mathematics questions. The variability of estimates of percentages of students having certain background characteristics or answering a certain cognitive question correctly is accounted for by the first component alone.

In addition to providing estimates of percentages of students and their proficiency, this report also provides information about the uncertainty of each statistic. Because NAEP uses complex sampling procedures, conventional formulas for estimating sampling variability that assume simple random sampling are inappropriate and NAEP uses a jackknife replication procedure to estimate standard errors. The jackknife standard error provides a reasonable measure of uncertainty for any information about students that can be observed without error, but each student typically responds to so few items within any content area that the proficiency measurement for any single student would be imprecise. In this case, using plausible values technology makes it possible to describe the performance of groups and subgroups of students, but the underlying imprecision that makes this step necessary adds an additional component of variability to statistics based on NAEP proficiencies.[19]

The reader is reminded that, like those from all surveys, NAEP results are also subject to other kinds of errors including the effects of necessarily imperfect adjustment for student and school nonresponse and other largely unknowable effects associated with the particular instrumentation and data collection methods used. Nonsampling errors can be attributed to a number of sources: inability to obtain complete information about all selected students in all selected schools in the sample (some students or schools refused to participate, or students participated but answered only certain items); ambiguous definitions; differences in interpreting questions; inability or unwillingness to give correct information; mistakes in recording, coding, or scoring data; and other errors of collecting, processing, sampling, and estimating missing data. The extent of nonsampling errors is difficult to estimate. By their nature, the impacts of such error cannot be reflected in the data-based estimates of uncertainty provided in NAEP reports.

---

[19] For further details, see Eugene G. Johnson, "Considerations and Techniques for the Analysis of NAEP Data" in *Journal of Educational Statistics* (December 1989).

Drawing Inferences from the Results

The use of *confidence intervals*, based on the standard errors, provides a way to make inferences about the population means and proportions in a manner that reflects the uncertainty associated with the sample estimates. An estimated sample mean proficiency ± 2 standard errors represents a 95 percent confidence interval for the corresponding population quantity. This means that with approximately 95 percent certainty, the average performance of the entire population of interest is within ± 2 standard errors of the sample mean.

As an example, suppose that the average mathematics proficiency of students in a particular group was 256, with a standard error of 1.2. A 95 percent confidence interval for the population quantity would be as follows:

Mean ± 2 standard errors = 256 ± 2 • (1.2) = 256 ± 2.4 =

256 - 2.4 and 256 + 2.4 = 253.6, 258.4

Thus, one can conclude with 95 percent certainty that the average proficiency for the entire population of students in that group is between 253.6 and 258.4.

Similar confidence intervals can be constructed for percentages, provided that the percentages are not extremely large (greater than 90) or extremely small (less than 10). For extreme percentages, confidence intervals constructed in the above manner may not be appropriate. However, procedures for obtaining accurate confidence intervals are quite complicated. Thus, comparisons involving extreme percentages should be interpreted with this in mind.

To determine whether there is a real difference between the mean proficiency (or proportion of a certain attribute) for two groups in the population, one needs to obtain an estimate of the degree of uncertainty associated with the difference between the proficiency means or proportions of these groups for the sample. This estimate of the degree of uncertainty -- called the standard error of the difference between the groups -- is obtained by taking the square of each group's standard error, summing these squared standard errors, and then taking the square root of this sum.

Similar to the manner in which the standard error for an individual group mean or proportion is used, the standard error of the difference can be used to help determine whether differences between groups in the population are real. The difference between the mean proficiency or proportion of the two groups ± 2 standard errors of the difference represents an approximate 95 percent confidence interval. If the resulting interval includes zero, there is insufficient evidence to claim a real difference between groups in the population. If the interval does not contain zero, the difference between groups is statistically significant (different) at the .05 level.

The procedures described in this section, and the certainty ascribed to intervals (e.g., a 95 percent confidence interval) are based on statistical theory that assumes that only one confidence interval or test of statistical significance is being performed. When one considers sets of confidence intervals, like those for the average proficiency of all participating states and territories, statistical theory indicates that the certainty associated with the entire set of intervals is less than that attributable to each individual comparison from the set. If one wants to hold the certainty level for a specific set of comparisons at a particular level (e.g., .95), adjustments (called multiple-comparisons procedures) need to be made.

The standard errors for means and proportions reported by NAEP are statistics and subject to a certain degree of uncertainty. In certain cases, typically when the standard error is based on a small number of students or when the group of students is enrolled in a small number of schools, the amount of uncertainty associated with the standard errors may be quite large. Throughout this report, estimates of standard errors subject to a large degree of uncertainty are designated by the symbol "!". In such cases, the standard errors -- and any confidence intervals or significance tests involving these standard errors -- should be interpreted cautiously.

ACKNOWLEDGMENTS

This report is culmination of effort by hundreds of individuals who contributed their considerable expertise, energy, and creativity to NAEP's mathematics assessments, from developing and conducting the assessments to analyzing and reporting the results. Most importantly, NAEP is grateful to the students and school staff who made the assessments possible.

The mathematics assessments, including the Trial State Assessment Program, were truly a collaborative effort among staff from State Education Agencies, the National Center for Education Statistics (NCES), Educational Testing Service (ETS), Westat, and National Computer Systems (NCS). The program benefited from the contributions of hundreds of individuals at the state and local levels -- Governors, Chief State School Officers, State and District Test Directors, State Coordinators, and district administrators -- who tireless provided their wisdom, experience, and hard work.

The assessment was funded through NCES, in the Office of Educational Research and Improvement of the U.S. Department of Education. Emerson Elliott, Commissioner, provided consistent support and guidance. The staff -- particularly Gary Phillips, Eugene Owen, Stephen Gorman, Maureen Treacy, Andrew Kolstad, and Sheida White -- worked closely and collegially with ETS, Westat, and NCS staff and played a crucial role in all aspects of the program. The members of the National Assessment Governing Board (NAGB) and the NAGB staff provided advice and guidance throughout.

The Chief State School Officers managed the National Assessment Planning Project that resulted in the mathematics framework and objectives for the assessment. NAEP owes a debt of gratitude to the numerous panelists and consultants who worked so diligently on developing the assessment and providing a frame for interpreting the results, including those who helped create the objectives and develop the assessment instruments.

Under the NAEP contract to ETS, Ina Mullis served as the Project Director. Stephen Koffler managed test development and state services. The mathematics test development team led by Chancey Jones and Jeffrey Haberstroh, and including Gail Chapman, Jane Kupin, Patty Klag, Jean MacFarland, Jane Maroney, Pat Meier, Melissa White, and Beverly Whittington, deserves special credit for its dedication. NAEP also owes a debt of gratitude to members of The Mathematics Item Development Panel who worked so diligently on developing the assessment.

Jules Goodison managed the operational aspects with the strong support of John Olson, who also provided invaluable help in managing state services and reporting. Sampling and data collection activities were carried out by Westat under the direction of Renee Slobasky, Nancy Caldwell, and Keith Rust. Printing, distribution, scoring, and processing activities were conducted by NCS, under the supervision of John O'Neill, Judy Moyer, Diane Smrdel, Lavonne Mohn, Brad Thayer, and Mathilde Kennel.

Mathematics and state design, statistical, psychometric procedures and implementation were led by John Mazzeo and Frank Jenkins under the direction of Eugene Johnson. John Barone managed the data analysis and computerized reporting systems. Within these endeavors, Robert Patrick and David Freund designed and implemented the table generation system. David Freund and Edward Kulick performed the mathematics data analysis, assisted by Drew Bowker, Bruce Kaplan, Steve Isham, and Yim Fai Fong. John Dossey, Ina Mullis, and Chancey Jones wrote the text for the report. Many thanks are provided to the numerous reviewers, internal and external, who suggested improvements to successive drafts, including Kent Ashworth, John Mazzeo, Andrew Latham, Jeffrey Haberstroh, Gail Chapman, Patricia Klag, and Dorothy Morrell. Sharon Davis-Johnson, with the assistance of Karen Damiano, provided the excellent desk-top publishing skills essential to the project. Kent Ashworth, who also manages information services, coordinated the cover design and production of the report.

☆ U.S. GOVERNMENT PRINTING OFFICE: 1993—342-109/93103